```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.


OM nucleic - nucleic search, using sw model

Run on:         November 29, 2004, 08:42:07 ; Search time 227 Seconds
                                             (without alignments)
                                             8623.403 Million cell updates/sec

Title:          US-09-989-920-100
Perfect score:  2754
Sequence:       1 gccagaagcagcctcagctt..........aaaaataaaaagatgcggcc 2754

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters:      1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/1/ina/backfiles1.seq:*

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.

                                 SUMMARIES
                      %
Result              Query
  No.    Score      Match Length DB  ID                           Description
-------------------------------------------------------------------------------
     1  1747.6   63.5    1853  3  US-09-232-160-3           Sequence 3, Appli
     2      52    1.9    7218  1  US-08-232-463-14          Sequence 14, Appl
     3      49    1.8    6330  3  US-09-880-427-2           Sequence 2, Appli
     4      49    1.8    6330  3  US-09-306-538B-2          Sequence 2, Appli
     5      49    1.8   19806  4  US-09-740-028A-3          Sequence 3, Appli
     6      49    1.8   19806  4  US-10-118-037-3           Sequence 3, Appli
     7      47    1.7     265  4  US-09-513-999C-27278      Sequence 27278, A
c    8    42.6    1.5  640681  4  US-09-790-988-1           Sequence 1, Appli
     9    41.6    1.5     210  4  US-09-248-796A-12137      Sequence 12137, A
c   10    41.4    1.5     329  4  US-09-513-999C-10733      Sequence 10733, A
c   11    40.8    1.5     318  4  US-09-513-999C-11322      Sequence 11322, A
    12    40.8    1.5     398  4  US-09-513-999C-23077      Sequence 23077, A
    13    40.4    1.5    1141  4  US-09-806-708B-22         Sequence 22, Appl
    14    40.4    1.5  640681  4  US-09-790-988-1           Sequence 1, Appli
c   15    39.6    1.4    1107  4  US-09-248-796A-8906       Sequence 8906, Ap
c   16    39.4    1.4     588  3  US-09-328-111-321         Sequence 321, App
    17    39.4    1.4    5852  1  US-07-867-106-2           Sequence 2, Appli
    18      39    1.4     303  4  US-09-601-198-179         Sequence 179, App
c   19    38.8    1.4    9636  1  US-08-323-170B-1          Sequence 1, Appli
c   20    38.8    1.4    9636  3  US-08-954-441-1           Sequence 1, Appli
c   21    38.4    1.4    2422  1  US-07-867-106-5           Sequence 5, Appli
c   22    38.4    1.4    5852  1  US-07-867-106-2           Sequence 2, Appli
c   23    38.4    1.4   54945  4  US-09-967-669-10          Sequence 10, Appl
c   24    38.2    1.4    1099  4  US-09-270-767-11371       Sequence 11371, A
c   25    38.2    1.4    3009  4  US-09-248-796A-6806       Sequence 6806, Ap
c   26    38.2    1.4    3454  4  US-09-270-767-11909       Sequence 11909, A
c   27    38.2    1.4    5185  3  US-08-971-395-5           Sequence 5, Appli
c   28    38.2    1.4    8561  3  US-09-112-450-3           Sequence 3, Appli
c   29    38.2    1.4    8561  4  US-09-419-291A-3          Sequence 3, Appli
    30    38.2    1.4   98844  4  US-09-791-211-10          Sequence 10, Appl
c   31      38    1.4    2007  2  US-08-743-637B-169        Sequence 169, App
c   32      38    1.4    2007  3  US-08-526-840B-169        Sequence 169, App
c   33      38    1.4    2028  3  US-09-134-001C-1710       Sequence 1710, Ap
c   34    37.8    1.4    4140  3  US-08-894-731-2           Sequence 2, Appli
c   35    37.8    1.4    8920  2  US-08-446-855A-1          Sequence 1, Appli
c   36    37.8    1.4    8920  3  US-09-150-741-1           Sequence 1, Appli
c   37    37.8    1.4 1664976  4  US-08-916-421B-1          Sequence 1, Appli
c   38    37.8    1.4 1664976  4  US-09-692-570-1           Sequence 1, Appli
c   39    37.4    1.4    2663  1  US-08-136-743B-3          Sequence 3, Appli
c   40    37.4    1.4    6040  4  US-10-204-708-70          Sequence 70, Appl
    41    37.4    1.4   74962  4  US-09-685-853A-3          Sequence 3, Appli
c   42    37.2    1.4     942  4  US-09-270-767-3767        Sequence 3767, Ap
c   43    37.2    1.4     942  4  US-09-270-767-19049       Sequence 19049, A
    44    37.2    1.4    1928  4  US-09-674-826B-5          Sequence 5, Appli
c   45    36.8    1.3   80246  3  US-09-078-294-4           Sequence 4, Appli


                                 ALIGNMENTS


RESULT 1
US-09-232-160-3
; Sequence 3, Application US/09232160
; Patent No. 6368794
; GENERAL INFORMATION:
;  APPLICANT: Steve Daniel
;  APPLICANT: James Gilmore
;  APPLICANT: Susan G. Stuart
;  APPLICANT: Laura Stuve
;  TITLE OF INVENTION: DETECTION OF ALTERED EXPRESSION OF GENES REGULATING CELL
;  TITLE OF INVENTION: PROLIFERATION
;  FILE REFERENCE: PA-0003 US
;  CURRENT APPLICATION NUMBER: US/09/232,160
;  CURRENT FILING DATE: 1999-01-15
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PERL Program
; SEQ ID NO 3
;   LENGTH: 1853
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: 1283330
US-09-232-160-3

  Query Match             63.5%;  Score 1747.6;  DB 3;  Length 1853;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1750;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        994 TGTTTTCTCCCCCATCACAAAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAA 1053
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TGTTTTCTCCCCCATCACAAAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAA 60

Qy       1054 AAATATGTACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGT 1113
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AAATATGTACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGT 120

Qy       1114 TTGCTCTTACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTT 1173
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTGCTCTTACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTT 180

Qy       1174 ACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCC 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCC 240

Qy       1234 ATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTTTTAAAAGAATAAA 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTTTTAAAAGAATAAA 300

Qy       1294 ATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAAC 1353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 ATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAAC 360

Qy       1354 CCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGG 1413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGG 420

Qy       1414 TGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTG 1473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTG 480

Qy       1474 CTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATC 1533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATC 540

Qy       1534 CCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGC 1593
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGC 600

Qy       1594 AGTGGCCTCCTGGCCAGCTCACACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCA 1653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGTGGCCTCCTGGCCAGCTCACACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCA 660

Qy       1654 CCTAGTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCA 1713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCTAGTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCA 720

Qy       1714 TGTGTCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGG 1773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGTGTCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGG 780

Qy       1774 TTTTGAGGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATC 1833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTTGAGGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATC 840

Qy       1834 TGCCATTGGAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGT 1893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCCATTGGAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGT 900

Qy       1894 CCGGCTGTGGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTT 1953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCGGCTGTGGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTT 960

Qy       1954 CACTCCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATA 2013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CACTCCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATA 1020

Qy       2014 GCAGCTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGA 2073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GCAGCTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGA 1080

Qy       2074 TCCTCCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAG 2133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TCCTCCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAG 1140

Qy       2134 GTTAGATGGGGGCCAGGGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGG 2193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GTTAGATGGGGGCCAGGGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGG 1200

Qy       2194 TGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTT 2253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTT 1260

Qy       2254 TGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAA 2313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAA 1320

Qy       2314 AAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTC 2373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTC 1380

Qy       2374 TCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTAT 2433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTAT 1440

Qy       2434 TTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCT 2493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCT 1500

Qy       2494 TTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGG 2553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGG 1560

Qy       2554 TGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTG 2613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTG 1620

Qy       2614 TACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTG 2673
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTG 1680

Qy       2674 TATGTGGATCTGTGCTTGGGTTAGAATGCAAATAAAACTCACATTTGTAAGAAAAAAAAA 2733
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       1681 TATGTGGATCTGTGCTTGGGTTAGAATGCAAATAAAACTCACATTTGTAAGACAAAAAAA 1740

Qy       2734 AAAAAATAAAAAGA 2747
              |||||  ||| |||
Db       1741 AAAAACAAAAGAGA 1754

RESULT 2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
```

```
US-08-232-463-14

  Query Match             1.9%;  Score 52;  DB 1;  Length 7218;
  Best Local Similarity   12.2%;  Pred. No. 0.00018;
  Matches   61;  Conservative  213;  Mismatches  228;  Indels    0;  Gaps    0;

Qy       1276 CTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGATG 1335
              | |||   | |   | | |  | |   |  | ||||||  |||||   ||    |    |
Db       1005 CGTTTGCCATACGCTCACAGAATTAATTCCGAGCTTGGCTGCAGGTCGAGGGAGCTTGCG 1064

Qy       1336 AGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGCC 1395
              | :  ::::: : ::  :::: ::: :: : : :::: : :::: : :  ::  :   ::
Db       1065 ATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1124

Qy       1396 CCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGTG 1455
              ::::::      ::::  : ::::  ::  ::::: :    ::: :::: :::: :  :
Db       1125 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1184

Qy       1456 GGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGCT 1515
                  : :    :::::: ::::::::::: :: ::: :::::: :::::     ::: ::
Db       1185 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1244

Qy       1516 GGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACTT 1575
                 :::::  ::: : ::::::  ::   ::::::    ::  : : :         :::
Db       1245 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1304

Qy       1576 TGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCTTGTCCTGGGAGGGGC 1635
              :   :  :  :   :  :  :  ::::::  ::  ::: : ::::: ::::        :
Db       1305 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1364

Qy       1636 AGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAG 1695
                :::  :::: ::::: :::  : ::::     ::     :::::: :::::::
Db       1365 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1424

Qy       1696 CCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACA 1755
              ::: :   ::    |  ||   || | ||  |   |    | || ||     |  | |||
Db       1425 YYYYYYYYYYGTACCAAATTCTTCTATCTCTTTAACTACTTGCATAGATAGGTAATTACA 1484

Qy       1756 GTGGCCTCAGCTTGGTGGTTTT 1777
              |||    |  | ||  | ||||
Db       1485 GTGATGCCTACATGCCGTTTTT 1506


RESULT 3
US-09-880-427-2
; Sequence 2, Application US/09880427
; Patent No. 6358728
; GENERAL INFORMATION:
;  APPLICANT: Simon, Andras
;  APPLICANT:  Eriksson, Ulf
;  APPLICANT:  Dryja, Thaddeus P.
;  APPLICANT:  Berson, Eliot
;  APPLICANT:  Yamamoto, Hioyuji
;  TITLE OF INVENTION: Mutations in Nucleic Acid Molecules Encoding 11-Cis Retinol
;  TITLE OF INVENTION:  Dehydrogenase, The Mutated Proteins, and Uses Thereof
;  FILE REFERENCE: LUD 5601
;  CURRENT APPLICATION NUMBER: US/09/880,427
;  CURRENT FILING DATE:  2001-06-13
;  PRIOR APPLICATION NUMBER: US 09/306,538
;  PRIOR FILING DATE: 1999-05-06
;  NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 2
;   LENGTH: 6330
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 5357, 5448
;   OTHER INFORMATION: nucleotide not determined
US-09-880-427-2

  Query Match             1.8%;  Score 49;  DB 3;  Length 6330;
  Best Local Similarity   56.5%;  Pred. No. 0.0013;
  Matches   91;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy       1061 TACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCT 1120
              |||    | ||  |  |||   || | |  |  |    |   || ||| |  | |
Db       4714 TACAAGCTGATAATGGTTTTTTTTATTTTTAAATGGTTACATTGTAAACTGTTATATAAG 4773

Qy       1121 TACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCT 1180
              |||   |  |||  ||| ||||||  |||||||     | ||  |||||||| ||  |||
Db       4774 TACCTGATAATATCATTAATTTTGTTTCTTGGCCTGCCATGCTTAAAATATTAACTCTCT 4833

Qy       1181 AGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
               |||| |||   |||||   |||||||  |||  | ||| |
Db       4834 GGCCCTTTAAGAAAAAAACGTGCTGACCCCTGCTCTAGATC 4874


RESULT 4
US-09-306-538B-2
; Sequence 2, Application US/09306538B
; Patent No. 6372463
; GENERAL INFORMATION:
;  APPLICANT: Simon, Andras
;  APPLICANT:  Eriksson, Ulf
;  APPLICANT:  Dryja, Thaddeus P.
;  APPLICANT:  Berson, Eliot
;  APPLICANT:  Yamamoto, Hioyuji
;  TITLE OF INVENTION: Mutations in Nucleic Acid Molecules Encoding 11-Cis Retinol
;  TITLE OF INVENTION:  Dehydrogenase, The Mutated Proteins, and Uses Thereof
;  FILE REFERENCE: LUD 5601
;  CURRENT APPLICATION NUMBER: US/09/306,538B
;  CURRENT FILING DATE:  1999-05-06
;  NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 2
;   LENGTH: 6330
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 5357, 5448
;   OTHER INFORMATION: nucleotide not determined
US-09-306-538B-2

  Query Match             1.8%;  Score 49;  DB 3;  Length 6330;
  Best Local Similarity   56.5%;  Pred. No. 0.0013;
  Matches   91;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy       1061 TACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCT 1120
              |||    | ||  |  |||   || | |  |  |    |   || ||| |  | |
Db       4714 TACAAGCTGATAATGGTTTTTTTTATTTTTAAATGGTTACATTGTAAACTGTTATATAAG 4773

Qy       1121 TACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCT 1180
              |||   |  |||  ||| ||||||  |||||||     | ||  |||||||| ||  |||
Db       4774 TACCTGATAATATCATTAATTTTGTTTCTTGGCCTGCCATGCTTAAAATATTAACTCTCT 4833

Qy       1181 AGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
               |||| |||   |||||   |||||||  |||  | ||| |
Db       4834 GGCCCTTTAAGAAAAAAACGTGCTGACCCCTGCTCTAGATC 4874


RESULT 5
US-09-740-028A-3
; Sequence 3, Application US/09740028A
; Patent No. 6410289
; GENERAL INFORMATION:
;  APPLICANT: GONG, Fangcheng et al
;  TITLE OF INVENTION: ISOLATED HUMAN DEHYDROGENASES, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING THESE HUMAN DEHYDROGENASES, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001054
;  CURRENT APPLICATION NUMBER: US/09/740,028A
;  CURRENT FILING DATE:  2000-12-20
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
```

```
    LENGTH: 19806
    TYPE: DNA
    ORGANISM: Human
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (1)...(19806)
    OTHER INFORMATION: n = A,T,C or G
US-09-740-028A-3

  Query Match             1.8%;  Score 49;  DB 4;  Length 19806;
  Best Local Similarity   56.5%;  Pred. No. 0.0031;
  Matches   91;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy       1061 TACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCT 1120
              |||    | ||  |  |||   || | |  |  |    |   || ||| |  | |
Db      16677 TACAAGCTGATAATGGTTTTTTTTATTTTTAAATGGTTACATTGTAAACTGTTATATAAG 16736

Qy       1121 TACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCT 1180
              |||   |  |||  ||| ||||||  |||||||     | ||  |||||||| ||  |||
Db      16737 TACCTGATAATATCATTAATTTTGTTTCTTGGCCTGCCATGCTTAAAATATTAACTCTCT 16796

Qy       1181 AGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
               |||| |||   |||||   |||||||  |||  | ||| |
Db      16797 GGCCCTTTAAGAAAAAAACGTGCTGACCCCTGCTCTAGATC 16837


RESULT 6
US-10-118-037-3
  Sequence 3, Application US/10118037
  Patent No. 6797499
  GENERAL INFORMATION:
   APPLICANT: GONG, Fangcheng et al.
   TITLE OF INVENTION: ISOLATED HUMAN DEHYDROGENASES, NUCLEIC
   TITLE OF INVENTION:  ACID MOLECULES ENCODING THESE HUMAN DEHYDROGENASES, AND USES
   TITLE OF INVENTION:  THEREOF
   FILE REFERENCE: CL001054DIV
   CURRENT APPLICATION NUMBER: US/10/118,037
   CURRENT FILING DATE:  2002-04-09
   PRIOR APPLICATION NUMBER: 09/740,028
   PRIOR FILING DATE: 2000-12-20
   NUMBER OF SEQ ID NOS: 4
   SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 3
    LENGTH: 19806
    TYPE: DNA
    ORGANISM: Homo sapien
    FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (1)...(19806)
    OTHER INFORMATION: n = A,T,C or G
US-10-118-037-3

  Query Match             1.8%;  Score 49;  DB 4;  Length 19806;
  Best Local Similarity   56.5%;  Pred. No. 0.0031;
  Matches   91;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy       1061 TACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCT 1120
              |||    | ||  |  |||   || | |  |  |    |   || ||| |  | |
Db      16677 TACAAGCTGATAATGGTTTTTTTTATTTTTAAATGGTTACATTGTAAACTGTTATATAAG 16736

Qy       1121 TACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCT 1180
              |||   |  |||  ||| ||||||  |||||||     | ||  |||||||| ||  |||
Db      16737 TACCTGATAATATCATTAATTTTGTTTCTTGGCCTGCCATGCTTAAAATATTAACTCTCT 16796

Qy       1181 AGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
               |||| |||   |||||   |||||||  |||  | ||| |
Db      16797 GGCCCTTTAAGAAAAAAACGTGCTGACCCCTGCTCTAGATC 16837


RESULT 7
US-09-513-999C-27278
  Sequence 27278, Application US/09513999C
  Patent No. 6783961
```

```
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 27278
;   LENGTH: 265
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 63
;   OTHER INFORMATION: k=g or t
US-09-513-999C-27278

  Query Match             1.7%;  Score 47;  DB 4;  Length 265;
  Best Local Similarity   62.2%;  Pred. No. 0.0005;
  Matches   74;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy       1103 TGGAAATTTGTTTGCTCTTACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGC 1162
              || ||| |  | |     |||   |  |||  ||| ||||||  |||||||     | ||
Db         20 TGTAAACTGTTATATAAGTACCTGATAATATCATTAATTTTGTKTCTTGGCCTGCCATGC 79

Qy       1163 CCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
                |||||||| ||  ||| |||| |||   |||||   |||||||  |||  | ||| |
Db         80 TTAAAATATTAACTCTCTGGCCCTTTAAGAAAAAAACGTGCTGACCCCTGCTCTAGATC 138


RESULT 8
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
;  APPLICANT: SHIGENOBU, SHUJI
;  APPLICANT:  WATANABE, HIDEMI
;  APPLICANT:  HATTORI, MASAHIRA
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;  FILE REFERENCE: 081356/0159
;  CURRENT APPLICATION NUMBER: US/09/790,988
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: JP2000-107160
;  PRIOR FILING DATE: 2000-04-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

  Query Match             1.5%;  Score 42.6;  DB 4;  Length 640681;
  Best Local Similarity   59.5%;  Pred. No. 3.7;
  Matches   72;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy       1023 TTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTGGA 1082
              || ||||||||  ||  ||| | |||  ||| || ||| |    |||||| | | | | |
Db     390713 TTTTTATTTTTTTAATCATATTTTTTTCAAGAAAATGAATGTCATTATTTTTTATGGAA 390654

Qy       1083 TTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGATTT 1142
               |||     ||||  || |  || || |||  |       ||    || ||  | |||||||
Db     390653 ATTTGTGTTTATTAAATCTGTAATTTTTCCCTGTCTATAAAAATATTAGTTTGATTT 390594

Qy       1143 T 1143
              |
```

```
Db     390593 T 390593

RESULT 9
US-09-248-796A-12137
; Sequence 12137, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 12137
;   LENGTH: 210
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-12137

  Query Match             1.5%;  Score 41.6;  DB 4;  Length 210;
  Best Local Similarity   67.0%;  Pred. No. 0.018;
  Matches   59;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy       1011 CAAAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTA 1070
              || ||||||| | ||  ||||||| |  |||||  ||||  ||||| |    ||||| ||
Db         32 CAGAAAAAAATAATCAAATTTTTAATCCACATGGTTTTAAGAAAAAGAAAATCTCAAATA 91

Qy       1071 TTGTATTTTGGATTTTATCAATTTAAAA 1098
              ||| ||||    ||| | | | | || |
Db         92 TTGCATTTCTTCTTTCAACTAATAAACA 119

RESULT 10
US-09-513-999C-10733/c
; Sequence 10733, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 10733
;   LENGTH: 329
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-513-999C-10733

  Query Match             1.5%;  Score 41.4;  DB 4;  Length 329;
  Best Local Similarity   62.6%;  Pred. No. 0.03;
  Matches   82;  Conservative    0;  Mismatches   46;  Indels    3;  Gaps    1;

Qy       1044 TATTTACCAAAAATATGTACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAATTGT 1103
              | | |  || ||| |  |     |||| | ||||||| |||||||||||  || |  |||
Db        323 TGTATCACACAAAAAATTCAAGTATTAATATATTTTGAATTTTATCAATAAAATATATGT 264

Qy       1104 GGAAATTTG---TTTGCTCTTACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAA 1160
              | ||||||    ||  |  | |    |  | |   |  ||||||||| |||||||  |||
Db        263 GAAAATTTTTCATTCTCATTAAATATATAAAAGCCTCAATTTTGCCTTTTGGCTCACAAA 204

Qy       1161 GCCCAAAATAT 1171
              ||| ||| | |
Db        203 GCCTAAACTTT 193

RESULT 11
US-09-513-999C-11322/c
; Sequence 11322, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 11322
;   LENGTH: 318
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 16
;   OTHER INFORMATION: r=a or g
US-09-513-999C-11322

  Query Match             1.5%;  Score 40.8;  DB 4;  Length 318;
  Best Local Similarity   52.3%;  Pred. No. 0.044;
  Matches   90;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy        750 AGTGGGAACCTGTCCTCATGCCCTTCGTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCC 809
              | || ||| ||   |    ||    || | |  || | ||  |||  | |||  |||||
Db        201 ACTGAGAAGCTCGGCCGCAGCAAGGCGGCGTCGGGGGGCCGCAGGGCGCCAGTGGGGGCA 142

Qy        810 AGGGGAAGTTTTCAGGCCTTCATCAAAGAGAACAACATCCTCAGCTCCGCACCCCTCATC 869
                ||  | |  ||||| |  |  || |   | |  ||||  || ||||||  ||||| ||
Db        141 GAGGCTATTCCTCAGGACACCGGCAGATCCATCCTCATCTCCATCTCCGCCACCCTCCTC 82

Qy        870 CTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTCAGCTAGCATACGGTG 921
              ||   |   | |       |      || |  ||||||| | | ||  | ||
Db         81 CTTCGTTGCCTCCCGGGCTTCCTCCTCTCCTGTTGTCAGTTGGGATCAGCTG 30

RESULT 12
US-09-513-999C-23077
; Sequence 23077, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 23077
;   LENGTH: 398
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 144
```

```
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 145
;   OTHER INFORMATION: m=a or c
US-09-513-999C-23077

  Query Match             1.5%;  Score 40.8;  DB 4;  Length 398;
  Best Local Similarity   63.0%;  Pred. No. 0.052;
  Matches   63;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy       1131 TAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTAC 1190
              || |  |||     |||||  |  |  ||||   |||||||||||  ||| ||   ||||
Db        197 TAGTTATGACAGAACCTCTATGGCCACAAAGTTTAAAATATTTACTATCTTGCTATTTAC 256

Qy       1191 AGAAAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACC 1230
              ||||||  | |||||||| |||  | | |    ||| | |
Db        257 AGAAAATATTTGCTGACTTCTGCCCTACAGAAACATAAGC 296


RESULT 13
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
;  FILE REFERENCE: 4810-58741
;  CURRENT APPLICATION NUMBER: US/09/806,708B
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/147,133
;  PRIOR FILING DATE: 1999-08-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 1141
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (1)..(1141)
;   OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

  Query Match             1.5%;  Score 40.4;  DB 4;  Length 1141;
  Best Local Similarity   9.4%;  Pred. No. 0.15;
  Matches   35;  Conservative  167;  Mismatches  169;  Indels    0;  Gaps    0;

Qy        965 AGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAAAATT 1024
              ::::: : : : : ::::::      : |: ::::::|     :  :|::: ::  | :
Db         53 MSKSRKWTWARMYCKYRRWYNNKSRWWKGWYKKKWYBCANNTSBRYHARRWKDMKTAYBM 112

Qy       1025 CTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTGGATT 1084
                | :: :|   ::  ::: :::|: :: : |    :   |::::::::|::::::
Db        113 TMTNKWGKTGWRHRYWRWRAMBDTVDHHYVTAMNNAWTTMCMMDKDDKRTRWWWKKNNNA 172

Qy       1085 TTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGATTTTG 1144
              | :: : :::::     :|: ::: :: : : ::    :|::::  :  ::|: :: :::
Db        173 TGWDDDTKYHMWNNNGCBTVTWMVRYKTDRDWSBKRMNYGMBWWKNWSYDVTYYWWVWDD 232

Qy       1145 CCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTCTGCT 1204
              :|: :  ::   |: ::: :::|: ||:  ::    |  : :: : :      |: : :
Db        233 MCKRKVRRWVRTRGRMRNYMVAWBTAHRRRYNNGWTBAMAYRRWTMNNNNNNAKAMCKRA 292

Qy       1205 GACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCA 1264
              :   :    : :     ::  |::   | |  |  |:  :::::::: ||  :
Db        293 KYWGWNRABVNSTCTTWKSKTTKVRTSCWANNCRAGDANKDHKWWKWSAAMGVYWNNNNN 352

Qy       1265 GACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAA 1324
                   : |: :  ::: :: |:: ::| ||:   ::  ::::  :        : :: :|
Db        353 NNWTYKKARHBARWDWVWHSAWKKWHANAAHYSRKKWTBYKRKTMVNNNNGTTMWKRMWA 412

Qy       1325 GCATTCAGATG 1335
```

```
                  :::  ::::|
Db        413 WYWKMDMDWBG 423


RESULT 14
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
;  APPLICANT: SHIGENOBU, SHUJI
;  APPLICANT:  WATANABE, HIDEMI
;  APPLICANT:  HATTORI, MASAHIRA
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;  FILE REFERENCE: 081356/0159
;  CURRENT APPLICATION NUMBER: US/09/790,988
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: JP2000-107160
;  PRIOR FILING DATE: 2000-04-07
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

  Query Match             1.5%;  Score 40.4;  DB 4;  Length 640681;
  Best Local Similarity   53.9%;  Pred. No. 17;
  Matches   83;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy        980 TGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAAAATTCTTATTTTTAGTAGA 1039
              | ||   ||| ||   | |  |   | || | | ||    |  |  | | |   | || |
Db     473336 TTAAAAAAATCAATATTATCTTTTTCAATTATATAATTCTATTTATTCAATAGAATTATA 473395

Qy       1040 CATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAA 1099
               |    | | || ||| |||| | | | ||| |   | ||   ||||   ||| ||| |||
Db     473396 TAAAAGTGTTCCTAAATTATGAAATAATTTTTAAAAATTAACATTTTAACAAATTAGAAA 473455

Qy       1100 TTGTGGAAATTTGTTTGCTCTTACGCCAACATAA 1133
              |     || | |||||| ||||  |   | ||||
Db     473456 TCAGTAAATTATGTTTGTTCTTTTGATGAAATAA 473489


RESULT 15
US-09-248-796A-8906/c
; Sequence 8906, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
;  APPLICANT: Keith Weinstock et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
;  TITLE OF INVENTION:  FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.132
;  CURRENT APPLICATION NUMBER: US/09/248,796A
;  CURRENT FILING DATE:  1999-02-12
;  PRIOR APPLICATION NUMBER: US 60/074,725
;  PRIOR FILING DATE: 1998-02-13
;  PRIOR APPLICATION NUMBER: US 60/096,409
;  PRIOR FILING DATE: 1998-08-13
;  NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 8906
;   LENGTH: 1107
;   TYPE: DNA
;   ORGANISM: Candida albicans
US-09-248-796A-8906

  Query Match             1.4%;  Score 39.6;  DB 4;  Length 1107;
  Best Local Similarity   52.4%;  Pred. No. 0.26;
  Matches   87;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy        975 ATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAAAATTCTTATTTTTA 1034
              || | ||||   |    | |||| | ||    || |  |  |  |  ||| ||||| |||
Db        218 ATTTCTGAATTTATATTACTGTTGTTTCTGTTATTATTATTATTATTATTATTATTATTA 159
```

```
Qy       1035 GTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTGGATTTTATCAATTT 1094
               ||    | |  ||  ||||| |   || |  ||||   |  | ||  |  |   |||||
Db        158 TTATTATTATTATTTGCAAAAGTCATTATTGTATTAGAATTCTGTGATTGATTATAATTT 99

Qy       1095 AAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140
                 |||||  || |||  | |  |  ||      ||| || || ||
Db         98 TGTAATTGATGAGATTGATGTCTTTGTAATTGTTCATTATTTTTAT 53

Search completed: November 29, 2004, 14:19:17
Job time : 236 secs
```

```
                  GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          November 29, 2004, 08:42:07 ; Search time 8386 Seconds
                                          (without alignments)
                                          11966.971 Million cell updates/sec

Title:           US-09-989-920-100
Perfect score:   2754
Sequence:        1 gccagaagcagcctcagctt..........aaaaataaaaagatgcggcc 2754

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters:      65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        EST:*
                1:  gb_est1:*
                2:  gb_est2:*
                3:  gb_htc:*
                4:  gb_est3:*
                5:  gb_est4:*
                6:  gb_est5:*
                7:  gb_est6:*
                8:  gb_gss1:*
                9:  gb_gss2:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                    %
Result            Query
  No.     Score   Match Length DB  ID                      Description
  -----------------------------------------------------------------------
     1   2726.8   99.0   3408  3  BC020554                BC020554 Homo sapi
     2      692   25.1    747  6  CD512690                CD512690 AGENCOURT
c    3    680.4   24.7    766  6  CB242367                CB242367 UI-CF-FN0
c    4      655   23.8    685  5  BU684754                BU684754 UI-CF-EN1
     5      640   23.2    642  5  BQ953674                BQ953674 AGENCOURT
c    6    629.8   22.9    791  4  BI518071                BI518071 603041774
c    7      615   22.3    640  5  BQ045123                BQ045123 UI-CF-EN1
c    8      608   22.1    625  6  CB305549                CB305549 UI-CF-EN1
     9    604.2   21.9    611  6  CD723979                CD723979 oj28g08.y
c   10    563.4   20.5    577  1  AI928242                AI928242 wo95h01.x
    11    562.8   20.4    768  4  BM019092                BM019092 603647129
    12    560.2   20.3    577  5  BU678926                BU678926 UI-CF-DU1
c   13      546   19.8    547  2  AW271767                AW271767 xs18a11.x
c   14    532.4   19.3    535  2  AW173671                AW173671 xj10g08.x
c   15      526   19.1    538  2  AW195353                AW195353 xn37d01.x
c   16      520   18.9    525  6  CA944724                CA944724 UI-CF-FN0
c   17    513.8   18.7    518  8  BZ369653                BZ369653 UP_473-19
c   18    513.4   18.6    515  1  AI434765                AI434765 ti20b01.x
c   19    513.2   18.6    518  1  AI678967                AI678967 tu60b06.x
c   20      508   18.4    508  1  AI151375                AI151375 qc84a01.x
c   21      506   18.4    506  2  BE138747                BE138747 xw94d11.x
c   22      506   18.4    507  5  BQ012939                BQ012939 UI-1-BC1p
c   23      501   18.2    501  1  AI831850                AI831850 wi52f07.x
c   24    500.4   18.2    522  2  BF223573                BF223573 7q77g11.x
    25      481   17.5    493  5  BQ082334                BQ082334 K-EST0087
c   26    479.2   17.4    487  2  BF592110                BF592110 7o04g03.x
c   27    476.4   17.3    490  2  BF437146                BF437146 7p67a02.x
c   28      473   17.2    473  1  AI453406                AI453406 tj37h08.x
c   29    467.4   17.0    472  5  BX110666                BX110666 BX110666
c   30    465.4   16.9    467  1  AI421402                AI421402 tf25e12.x
c   31      456   16.6    468  1  AI347069                AI347069 qp60c01.x
c   32    454.4   16.5    456  2  AW874390                AW874390 hp99d01.x
c   33    447.2   16.2    485  2  BE857167                BE857167 7g21d12.x
c   34      447   16.2    449  1  AI245431                AI245431 qk29e02.x
c   35    445.2   16.2    454  1  AI733680                AI733680 ov13h01.x
c   36      445   16.2    445  1  AI241361                AI241361 qk17e10.x
c   37      443   16.1    443  1  AI189643                AI189643 qd18f08.x
c   38    441.2   16.0    446  2  AW772502                AW772502 hn76d05.x
    39    436.8   15.9    440  4  BM055156                BM055156 ie91b10.y
c   40    434.4   15.8    436  1  AI522077                AI522077 ti78d10.x
c   41      429   15.6    429  1  AI347815                AI347815 qp02g08.x
c   42      425   15.4    425  1  AI985135                AI985135 wr98b04.x
c   43    424.2   15.4    460  2  BE671688                BE671688 7b60e04.x
c   44    423.4   15.4    425  2  AW439823                AW439823 hb88d02.x
c   45      420   15.3    445  1  AI985235                AI985235 ws01a03.x

                               ALIGNMENTS

RESULT 1
BC020554
LOCUS       BC020554                3408 bp    mRNA    linear   HTC 04-MAR-2003
DEFINITION  Homo sapiens, clone IMAGE:3049181, mRNA.
ACCESSION   BC020554
VERSION     BC020554.1  GI:18088242
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3408)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:      http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 20 Row: a Column: 5
            This clone has the following problem: retained intron.
FEATURES             Location/Qualifiers
     source          1. .3408
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3049181"
                     /tissue_type="Placenta, choriocarcinoma"
                     /clone_lib="NIH_MGC_10"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN
```

```
Query Match            99.0%;  Score 2726.8;  DB 3;  Length 3408;
Best Local Similarity   99.7%;  Pred. No. 0;
Matches 2731;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        671 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 730

Qy         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 790

Qy        121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180
              |||||||||| |||||||||||||||||||||||||||||||| | ||||||||||| |
Db        791 GAGTGGATGAAGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAGTGGCACTGGGCGAGGC 850

Qy        181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        851 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 910

Qy        241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        911 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 970

Qy        301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 1030

Qy        361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1031 GGCCAGTGAGGTGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 1090

Qy        421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 1150

Qy        481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1151 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 1210

Qy        541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1211 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 1270

Qy        601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1271 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 1330

Qy        661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1331 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 1390

Qy        721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1391 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 1450

Qy        781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1451 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 1510

Qy        841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1511 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 1570

Qy        901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1571 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 1630

Qy        961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1631 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1690

Qy       1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1691 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1750

Qy       1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1751 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1810

Qy       1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1811 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1870

Qy       1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1871 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1930

Qy       1261 CTCAGACAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1931 CTCAGACAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1990

Qy       1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1991 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 2050

Qy       1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2051 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 2110

Qy       1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2111 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 2170

Qy       1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2171 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 2230

Qy       1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2231 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 2290

Qy       1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2291 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 2350

Qy       1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2351 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 2410

Qy       1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2411 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 2470

Qy       1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2471 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 2530

Qy       1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2531 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 2590

Qy       1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2591 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 2650

Qy       1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2651 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2710

Qy       2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2711 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2770

Qy       2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        2771 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2830

Qy        2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2831 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2890

Qy        2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2891 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2950

Qy        2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2951 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 3010

Qy        2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3011 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 3070

Qy        2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3071 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 3130

Qy        2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3131 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 3190

Qy        2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3191 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 3250

Qy        2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3251 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 3310

Qy        2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3311 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 3370

Qy        2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAA 2738
               |||||||||||||||||||||||| |||||||||||||
Db        3371 TGCAAATAAAACTCACATTTGTAAAAAAAAAAAAAAAA 3408


RESULT 2
CD512690
LOCUS       CD512690                 747 bp    mRNA    linear   EST 06-JUN-2003
DEFINITION  AGENCOURT_14360615 NIH_MGC_186 Homo sapiens cDNA clone
            IMAGE:30406456 5', mRNA sequence.
ACCESSION   CD512690
VERSION     CD512690.1  GI:31444408
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 747)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDCM195  row: d  column: 17
            High quality sequence stop: 509.
FEATURES             Location/Qualifiers
     source          1. .747
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30406456"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_186"
                     /note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site_1: SfiI
                     (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); Library is
                     oligo-dT primed and directionally cloned. cDNA was
                     prepared from a pooled samples of tissues from Skin,
                     meninges, duramatter, pia matter and choroid plexus.   5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C, G, or T). Average
                     insert size 1.47 kb (range 0.50-4.0 kb).  15/15 colonies
                     contained inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                     Library"
ORIGIN

  Query Match             25.1%;  Score 692;  DB 6;  Length 747;
  Best Local Similarity   98.5%;  Pred. No. 8.3e-171;
  Matches  709;  Conservative    0;  Mismatches   10;  Indels    1;  Gaps    1;

Qy        1842 GAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGT 1901
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 GGGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGT 60

Qy        1902 GGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAG 1961
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAG 120

Qy        1962 GGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGG 2021
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGG 180

Qy        2022 GGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCT 2081
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 GGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCT 240

Qy        2082 AAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATG 2141
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATG 300

Qy        2142 GGGGCCAGGGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGT 2201
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGGGCCAGGGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGT 360

Qy        2202 CCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAG 2261
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAG 420

Qy        2262 TGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTG 2321
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTG 480

Qy        2322 CGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGC 2381
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGC 540

Qy        2382 CTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATG 2441
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATG 600

Qy        2442 ATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAA 2501
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 ATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAA 660

Qy        2502 GGTGGTGGGAGTATCTGTCGGGGTGGTGTGG-CCCTTGGATGGGTCAGGTGGGTGTGAGA 2560
```

```
              || |||||||||||||||||||||||||||| |||||||||||||      || ||| ||
Db        661 GGGGGTGGAGTATCTGTCGGGGTGGTGTGGCCCCTTGGATGGGTTCAAGTGGGGTGTGA 720


RESULT 3
CB242367/c
LOCUS       CB242367                 766 bp    mRNA    linear   EST 12-FEB-2003
DEFINITION  UI-CF-FN0-age-m-22-0-UI.s1 UI-CF-FN0 Homo sapiens cDNA clone
            UI-CF-FN0-age-m-22-0-UI 3', mRNA sequence.
ACCESSION   CB242367
VERSION     CB242367.1  GI:28364011
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 766)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..766
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-FN0-age-m-22-0-UI"
                     /tissue_type="Human Lung Epithelial cells"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-FN0"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-FN0 is a subtracted cDNA library derived from two
                     normalized Human lung epithelial cell libraries (EN1 and
                     DU1) The library was subtracted according to according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. For additional information, contact:
                     bento-soares@uiowa.edu
                     TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                     6hr to LPS 24h
                     TAG_LIB=UI-CF-FN0
                     TAG_SEQ=CTGCTCAGGT"
ORIGIN

  Query Match             24.7%;  Score 680.4;  DB 6;  Length 766;
  Best Local Similarity   97.6%;  Pred. No. 9.6e-168;
  Matches  723;  Conservative    0;  Mismatches   12;  Indels    6;  Gaps    3;

Qy       1999 GTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGT 2058
              ||||||||||||||||||||||||      |         ||||||||||||||| | ||
Db        736 GTAACAGTAGTGATAGCAGCTGGG----CTAACAGCTAGCTTGTGTTCTGCGCATNT-GT 682

Qy       2059 CAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAG 2118
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        681 CAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGG-GGCCCCCACTAGCCCAGTTTTCAG 623
```

```
Qy       2119 GAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCCACAGCTACTGATGGCCCGAGCCA 2178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        622 GAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCCACAGCTACTGATGGCCCGAGCCA 563

Qy       2179 GGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGG 2238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        562 GGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGG 503

Qy       2239 TGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGT 2298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        502 TGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGT 443

Qy       2299 GTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGAC 2358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        442 GTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGAC 383

Qy       2359 TGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCT 2418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        382 TGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCT 323

Qy       2419 CTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCC 2478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        322 CTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCC 263

Qy       2479 AGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTG 2538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        262 AGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTG 203

Qy       2539 GATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTTGTCCT 2598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        202 GATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTTGTCCT 143

Qy       2599 ACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTG 2658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        142 ACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTG 83

Qy       2659 CGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAATGCAAATAAAACTCACAT 2718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 CGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAATGCAAATAAAACTCACAT 23

Qy       2719 TTGTAAGAAAAAAAAAAAAAA 2739
              |||||||||||||||||||||
Db         22 TTGTAAGAAAAAAAAAAAAAA 2


RESULT 4
BU684754/c
LOCUS       BU684754                 685 bp    mRNA    linear   EST 07-OCT-2002
DEFINITION  UI-CF-EN1-acv-d-09-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
            UI-CF-EN1-acv-d-09-0-UI 3', mRNA sequence.
ACCESSION   BU684754
VERSION     BU684754.1  GI:23538028
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
```

```
              cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
              cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-EN1-acv-d-09-0-UI"
                     /tissue_type="Primary Lung Cystic Fibrosis Epithelial
                     Cells"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-EN1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-EN1 is a normalized cDNA library containing the
                     following tissue(s): Primary Lung Cystic Fibrosis
                     Epithelial Cells. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site.  Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CTGCTCAGGT.
                     TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                     6hr to LPS 24h
                     TAG_LIB=UI-CF-EN1
                     TAG_SEQ=CTGCTCAGGT"
ORIGIN

  Query Match             23.8%;  Score 655;  DB 5;  Length 685;
  Best Local Similarity   98.5%;  Pred. No. 4.7e-161;
  Matches  669;  Conservative    0;  Mismatches    9;  Indels    1;  Gaps    1;

Qy        345 AGGGACAGGGCTTGAGGGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACC 404
              ||| ||||||||||||||||||||||| |||||||||||||| |||||||||||| ||||
Db        685 AGGNACAGGGCTTGAGGGCCAGTGAGGTGCTGGTCAGGCACCNCAGGCCTCCTTGNGACC 626

Qy        405 TGCCCAGGGGCACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAA 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 TGCCCAGGGGCACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAA 566

Qy        465 AAGCTCCTCCTGGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGG 524
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 AAGCTCCTCCTGGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGG 506

Qy        525 GCAACTGCTGTGCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGG 584
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 GCAACTGCTGTGCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGG 446

Qy        585 CCTTCTAGAGGCAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGC 644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 CCTTCTAGAGGCAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGC 386

Qy        645 AGGGCAGCTGATGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAG 704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 AGGGCAGCTGATGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAG 326

Qy        705 CCTGGCGATGGGGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCC 764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 CCTGGCGATGGGGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCC 266

Qy        765 TCATGCCCTTCGTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAG 824
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 TCATGCCCTTCGTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAG 206

Qy        825 GCCTTCATCAAAGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTA 884
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GCCTTCATCAAAGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTA 146

Qy        885 CCGGTGTGTGACTGCCCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGT 944
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 CCGGTGTGTGACTGCCCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGT 86

Qy        945 TTATGCCACTGATTTATGATAGGGAATATTATCTTTGAACCCAA-TGAAGTGTTTTCTCC 1003
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db         85 TTATGCCACTGATTTATGATAGGGAATATTATCTTTGAACCCAAGTGAAGTGTTTTCTCC 26

Qy       1004 CCCATCACAAAAAAAAAAA 1022
              ||||     ||||||||||
Db         25 CCCANTCACAAAAAAAAAA 7


RESULT 5
BQ953674
LOCUS       BQ953674                 642 bp    mRNA    linear   EST 21-AUG-2002
DEFINITION  AGENCOURT_8803100 Lupski_sciatic_nerve Homo sapiens cDNA clone
            IMAGE:6198871 5', mRNA sequence.
ACCESSION   BQ953674
VERSION     BQ953674.1  GI:22369152
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13611  row: d  column: 08
            High quality sequence stop: 600.
FEATURES             Location/Qualifiers
     source          1. .642
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6198871"
                     /sex="male"
                     /tissue_type="sciatic nerve"
                     /dev_stage="adult, 70 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_sciatic_nerve"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.87 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
ORIGIN

  Query Match             23.2%;  Score 640;  DB 5;  Length 642;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-157;
  Matches  640;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       2100 CCACTAGCCCAGTTTTCAGGAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCCACAG 2159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCACTAGCCCAGTTTTCAGGAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCCACAG 60

Qy       2160 CTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGA 2219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGA 120

Qy       2220 TCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAG 2279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAG 180

Qy       2280 GCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAA 2339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAA 240

Qy       2340 CCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGA 2399
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGA 300

Qy       2400 AGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGT 2459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGT 360

Qy       2460 AACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGT 2519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGT 420

Qy       2520 CGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCG 2579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCG 480

Qy       2580 TTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCAC 2639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCAC 540

Qy       2640 TTGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGA 2699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGA 600

Qy       2700 ATGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAA 2739
              ||||||||||||||||||||||||||||||||||||||||
Db        601 ATGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAA 640


RESULT 6
BI518071/c
LOCUS       BI518071                 791 bp    mRNA    linear   EST 29-AUG-2001
DEFINITION  603041774T1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5162851 3',
            mRNA sequence.
ACCESSION   BI518071
VERSION     BI518071.1  GI:15342863
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 791)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11404  row: d  column: 20
            High quality sequence start: 10
            High quality sequence stop: 759.
FEATURES             Location/Qualifiers
     source          1..791
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5162851"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_116"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female.  Library is
                     oligo-dT primed and  directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.4 kb,
                     insert size range 1-3 kb.  Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber  (Invitrogen).  Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             22.9%;  Score 629.8;  DB 4;  Length 791;
  Best Local Similarity   96.2%;  Pred. No. 2.2e-154;
  Matches  753;  Conservative    0;  Mismatches   17;  Indels   13;  Gaps   10;

Qy       1882 TGGGTCCGAGGTCCGGCTGTGGCCAT--CCAGCCCCCTGTGGC-TTGTCCAGCCTCTGTG 1938
              ||| ||||||||||||||||||||||  |  |||||||||||| ||||||||||||||||
Db        791 TGGTTCCGAGGTCCGGCTGTGGCCATTCCAGGCCCCCTGTGGCATTGTCCAGCCTCTGTG 732

Qy       1939 CACCCCTGGTGTCTTCACT-CCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTG 1997
              ||| ||||||||||||||| |||||||||||||||||||||||||| |||| ||||||||
Db        731 CACACCTGGTGTCTTCACTCCCAGGGGCAGACAGCAGCCACTGCAGCTCCTCTCTTCGTG 672

Qy       1998 AGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTTGTGTTC-TGCGCATTTG 2056
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        671 AGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTTGTGTTCATGCGCATTTG 612

Qy       2057 GTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTC 2116
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 GTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTC 552

Qy       2117 AGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGCTACTGATGGCCCGAGC 2176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        551 AGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGCTACTGATGGCCCGAGC 492

Qy       2177 CAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATA 2236
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 CAGGTCGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATA 432

Qy       2237 GGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGG 2296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 GGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGG 372

Qy       2297 GTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGC-TTGGTTAACCACAGAAGAGAAGGG 2355
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        371 GTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCATTGGTTAACCACAGAAGAGAAGGG 312

Qy       2356 GACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGG-TGGAAGCACGGT-TACTGT 2413
              |||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||||
Db        311 GACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGCTGGAAGCACGGTCTACTGT 252

Qy       2414 GTTCTCTAA---TGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTAACCTCCACAC 2470
              ||||||| |   |   |     ||||||||||||||||||||||||||||||||||||||
Db        251 GTTCTCTCAATGTTCACTGTACTTTAAAATGATTTCTTTCTAAAGATGTAACCTCCACAC 192

Qy       2471 CTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTG-TCGGGGTGGTG 2529
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        191 CTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGCTCGGGGTGGTG 132

Qy       2530 TGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCT-GGGGAGGTGGGCGTTGAGCTCA 2588
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        131 TGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTCGGGGAGGTGGGCGTTGAGCTCA 72
```

```
Qy       2589 AAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTTACTG 2648
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db         71 AAGTTGTCCTACTGCCATGTTTGTGTACCTGAAATAAAGCATATTTTGCACTTGTTACTG 12

Qy       2649 TAC 2651
              |||
Db         11 TAC 9


RESULT 7
BQ045123/c
LOCUS       BQ045123                 640 bp    mRNA    linear   EST 21-FEB-2003
DEFINITION  UI-CF-EN1-ael-i-10-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
            UI-CF-EN1-ael-i-10-0-UI 3', mRNA sequence.
ACCESSION   BQ045123
VERSION     BQ045123.1  GI:19796176
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 640)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..640
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-EN1-ael-i-10-0-UI"
                     /tissue_type="Primary Lung Cystic Fibrosis Epithelial
                     Cells"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-EN1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-EN1 is a normalized cDNA library containing the
                     following tissue(s): Primary Lung Cystic Fibrosis
                     Epithelial Cells. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site.  Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CTGCTCAGGT.
                     TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                     6hr to LPS 24h
                     TAG_LIB=UI-CF-EN1
                     TAG_SEQ=CTGCTCAGGT"
```

```
ORIGIN

  Query Match             22.3%;  Score 615;  DB 5;  Length 640;
  Best Local Similarity   98.6%;  Pred. No. 1.6e-150;
  Matches  630;  Conservative    0;  Mismatches    7;  Indels    2;  Gaps    1;

Qy       2103 CTAGCCCAGTTTTCAGG--AAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160
              ||||||||||| |||||  |    |  |||||||||||||||||||||||||||||||||
Db        640 CTAGCCCAGTTNTCAGGGAAGTCAACNTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 581

Qy       2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 521

Qy       2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 461

Qy       2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 401

Qy       2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 341

Qy       2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 281

Qy       2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 221

Qy       2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 161

Qy       2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 101

Qy       2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 41

Qy       2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAA 2739
              |||||||||||||||||||||||||||||||||||||||
Db         40 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAA 2


RESULT 8
CB305549/c
LOCUS       CB305549                 625 bp    mRNA    linear   EST 04-MAR-2003
DEFINITION  UI-CF-EN1-aed-n-21-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
            UI-CF-EN1-aed-n-21-0-UI 3', mRNA sequence.
ACCESSION   CB305549
VERSION     CB305549.1  GI:28846060
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 625)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
```

```
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..625
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-EN1-aed-n-21-0-UI"
                     /tissue_type="Primary Lung Cystic Fibrosis Epithelial
                     Cells"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-EN1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-EN1 is a normalized cDNA library containing the
                     following tissue(s): Primary Lung Cystic Fibrosis
                     Epithelial Cells. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site.  Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pT7T3-Pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is CTGCTCAGGT.
                     TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
                     6hr to LPS 24h
                     TAG_LIB=UI-CF-EN1
                     TAG_SEQ=CTGCTCAGGT"
ORIGIN

  Query Match            22.1%;  Score 608;  DB 6;  Length 625;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-148;
  Matches  619;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        404 CTGCCCAGGGGCACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCA 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 CTGCCCAGGGGCACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCA 566

Qy        464 AAAGCTCCTCCTGGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAG 523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 AAAGCTCCTCCTGGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAG 506

Qy        524 GGCAACTGCTGTGCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGG 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 GGCAACTGCTGTGCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGG 446

Qy        584 GCCTTCTAGAGGCAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTG 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 GCCTTCTAGAGGCAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTG 386

Qy        644 CAGGGCAGCTGATGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCA 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 CAGGGCAGCTGATGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCA 326

Qy        704 GCCTGGCGATGGGGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTC 763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 GCCTGGCGATGGGGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTC 266

Qy        764 CTCATGCCCTTCGTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCA 823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 CTCATGCCCTTCGTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCA 206

Qy        824 GGCCTTCATCAAAGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTT 883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 GGCCTTCATCAAAGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTT 146

Qy        884 ACC-GGTGTGTGACTGCCCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCT 942
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 ACCNGGTGTGTGACTGCCCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCT 86

Qy        943 GTTTATGCCACTGATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTC 1002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GTTTATGCCACTGATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTC 26

Qy       1003 CCCCATCACAAAAAAAAAAA 1022
              ||||||||||||||||||||
Db         25 CCCCATCACAAAAAAAAAAA 6


RESULT 9
CD723979
LOCUS       CD723979                 611 bp    mRNA    linear   EST 26-JUN-2003
DEFINITION  oj28g08.y1 Human lacrimal gland, unamplified: oj Homo sapiens cDNA
            clone oj28g08 5', mRNA sequence.
ACCESSION   CD723979
VERSION     CD723979.1  GI:32274833
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 611)
  AUTHORS   Dickinson,D., Laurie,G. and Wistow,G.
  TITLE     Expressed sequence tag analysis of human lacrimal gland
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 28  row: g  column: 08
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..611
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="oj28g08"
                     /tissue_type="lacrimal gland"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human lacrimal gland, unamplified: oj"
                     /note="Organ: Eye; Vector: pCMVSport6; RNA was extracted
                     from 2 human lacrimal glands. A directionally cloned cDNA
                     library in the pCMVSPORT6 vector(Life Technologies) was
                     constructed at Bioserve Biotechnology (Laurel MD)
                     essentially following the protocols of the SuperScript
                     Plasmid System full details of which are contained in the
                     manufacturer's Instruction manual
                     (http://www.lifetech.com/). First strand synthesis was
                     carried out using a Not I primer-adapter
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. EST analysis
                     was performed on the unamplified library at the NIH
                     Intramural Sequencing Center (NISC)."
ORIGIN

  Query Match            21.9%;  Score 604.2;  DB 6;  Length 611;
  Best Local Similarity   99.5%;  Pred. No. 1.1e-147;
  Matches  606;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy          167 GCACTGGGCGATGTCCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTG 226
                ||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Db            2 GCACTGGGCGAGGCCCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTG 61

Qy          227 GCTTCACACAGGCACTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGT 286
                |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db           62 GCTTCACACAGGCACTTCCGCTTGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGT 121

Qy          287 TGCAGGTAGGAAATGTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAG 346
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          122 TGCAGGTAGGAAATGTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAG 181

Qy          347 GGACAGGGCTTGAGGGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTG 406
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          182 GGACAGGGCTTGAGGGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTG 241

Qy          407 CCCAGGGGCACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAA 466
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          242 CCCAGGGGCACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAA 301

Qy          467 GCTCCTCCTGGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGC 526
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          302 GCTCCTCCTGGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGC 361

Qy          527 AACTGCTGTGCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCC 586
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          362 AACTGCTGTGCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCC 421

Qy          587 TTCTAGAGGCAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAG 646
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          422 TTCTAGAGGCAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAG 481

Qy          647 GGCAGCTGATGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCC 706
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          482 GGCAGCTGATGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCC 541

Qy          707 TGGCGATGGGGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTC 766
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          542 TGGCGATGGGGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTC 601

Qy          767 ATGCCCTTC 775
                |||||||||
Db          602 ATGCCCTTC 610


RESULT 10
AI928242/c
LOCUS       AI928242                 577 bp    mRNA    linear   EST 08-MAR-2000
DEFINITION  wo95h01.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2463121 3',
            mRNA sequence.
ACCESSION   AI928242
VERSION     AI928242.1  GI:5664206
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 721    Std Error: 0.00
```

```
            Seq primer: -40UP from Gibco
            High quality sequence stop: 453.
FEATURES             Location/Qualifiers
     source          1..577
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2463121"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid11"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "
ORIGIN

  Query Match             20.5%;  Score 563.4;  DB 1;  Length 577;
  Best Local Similarity   99.7%;  Pred. No. 6.5e-137;
  Matches  575;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy         2150 GGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCC 2209
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          577 GGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCC 518

Qy         2210 CACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCT 2269
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          517 CACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCT 458

Qy         2270 CTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTT 2329
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          457 CTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTT 398

Qy         2330 GCTTGGTTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGT 2389
                ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          397 GCTTGGTTAACCACAGAAGAGAAGGGGACTGTTTGGGGTACCTCTCTGCAGCCTCCCCGT 338

Qy         2390 GCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTT 2449
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          337 GCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTT 278

Qy         2450 CTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAA-GGTGGTG 2508
                |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db          277 CTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGGTGGTG 218

Qy         2509 GGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGG 2568
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          217 GGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGG 158

Qy         2569 GGAGGTGGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGC 2628
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          157 GGAGGTGGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGC 98

Qy         2629 ATATTTTGCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTG 2688
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           97 ATATTTTGCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTG 38

Qy         2689 CTTGGGTTAGAATGCAAATAAAACTCACATTTGTAAG 2725
                |||||||||||||||||||||||||||||||||||||
Db           37 CTTGGGTTAGAATGCAAATAAAACTCACATTTGTAAG 1


RESULT 11
BM019092
LOCUS       BM019092                 768 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603647129F1 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5428711 5',
            mRNA sequence.
ACCESSION   BM019092
```

```
VERSION     BM019092.1  GI:16533446
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 768)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1896  row: j  column: 08
            High quality sequence stop: 578.
FEATURES             Location/Qualifiers
     source          1..768
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5428711"
                     /tissue_type="astrocytoma grade IV, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_98"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             20.4%;  Score 562.8;  DB 4;  Length 768;
  Best Local Similarity   94.8%;  Pred. No. 1e-136;
  Matches  626;  Conservative    0;  Mismatches   27;  Indels    7;  Gaps    4;

Qy       1174 ACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCC 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 ACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCC 61

Qy       1234 ATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTTTTAAAAGAATAAA 1293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 ATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTTTTAAAAGAATAAA 121

Qy       1294 ATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAAC 1353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAAC 181

Qy       1354 CCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGG 1413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGG 241

Qy       1414 TGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTG 1473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTG 301

Qy       1474 CTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATC 1533
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        302 CTTTCCTTCTTGCCAGCTGCTTCCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATC 361

Qy       1534 CCCTGGCCAGGTCCCTCAGGG-TTGATGCGTGGAGAA-GGACTTTGAGCAGTGGTGGGCA 1591
              ||||||||||||||||||||| ||||||||||||||| ||||||||||||||||||||||
Db        362 CCCTGGCCAGGTCCCTCAGGGATTGATGCGTGGAGAACGGACTTTGAGCAGTGGTGGGCA 421

Qy       1592 GCAGTGGCCTCCTGGCCAGCTCACACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTC 1651
```

```
              ||||||||||||||||||| ||| ||||| |||||||||||||||||||||||||| |||
Db        422 GCAGTGGCCTCCTGGCCAGATCAAACTCTAGTCCTGGGAGGGGCAGCCTGATCTCAACTC 481

Qy       1652 CACCTAGTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCC 1711
              || ||||||||||||||||||||||||||| ||||||||||||||||||||| |||||||
Db        482 CAACTAGTACCTTGGGGACTGAGGACCTTTAGGCTTCTCTGGAGCCTGCAAGACTCTTCC 541

Qy       1712 CATGTGTCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCAC-AGTGGCCTCAGCTTGG 1770
              |||||||||||||||||||||||||||||||||||||| |||| ||||| |||||||||
Db        542 CATGTGTCCAGCTGCTCTTCCTGCTACAAAGGGGACTGATCACAAGTGGACTCAGCTTGA 601

Qy       1771 TGGTTTTGAGGGGCCGCC----CCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATT 1826
                |       | |  |||    | |||| ||||||||||||||||||||||||||||| |
Db        602 AAGGAAAAGAGAGGAGCCGACACACCGGACCTCCATAAGGGTATCCTGGGCCTGAGAACT 661


RESULT 12
BU678926
LOCUS       BU678926                 577 bp    mRNA    linear   EST 07-OCT-2002
DEFINITION  UI-CF-DU1-abc-n-17-0-UI.s1 UI-CF-DU1 Homo sapiens cDNA clone
            UI-CF-DU1-abc-n-17-0-UI 3', mRNA sequence.
ACCESSION   BU678926
VERSION     BU678926.1  GI:23526358
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu
            Tissue Procurement: Dr. M. J. Welsh, University of Iowa
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com) or from Open Biosystems
            (www.openbiosystems.com).
             The following repetitive elements were found in this cDNA
            sequence: 188-250, >MER58B#DNA/MER1_type (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..577
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-CF-DU1-abc-n-17-0-UI"
                     /tissue_type="Primary Lung Epithelial Cells"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                     /clone_lib="UI-CF-DU1"
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     UI-CF-DU1 is a normalized cDNA library containing the
                     following tissue(s): Primary Lung Epithelial Cells The
                     library was constructed according to Bonaldo, Lennon and
                     Soares, Genome Research, 6:791-806, 1996. First strand
                     cDNA synthesis was primed with an oligo-dT primer
                     containing a Not I site.  Double stranded cDNA was ligated
                     to an EcoR I adaptor, digested with Not I, and cloned
                     directionally into pT7T3-Pac vector. The oligonucleotide
```

```
            used to prime the synthesis of first-strand cDNA contains
            a library tag sequence that is located between the Not I
            site and the (dT)18 tail. The sequence tag for this
            library is GGCTGTAGGC.
            TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
            TAG_LIB=UI-CF-DU1
            TAG_SEQ=GGCTGTAGGC"
ORIGIN

  Query Match             20.3%;  Score 560.2;  DB 5;  Length 577;
  Best Local Similarity   99.5%;  Pred. No. 4.6e-136;
  Matches  562;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        971 TATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAAAAATTCTTATT 1030
              | |||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db          9 TTTTATCTTTGAACCCAATGACATGTTTTCTCCCCCATCACAAAAAAAAAAAATTCTTATT 68

Qy       1031 TTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTGGATTTTATCA 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 TTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTGGATTTTATCA 128

Qy       1091 ATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGATTTTGCCTCTT 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 ATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGATTTTGCCTCTT 188

Qy       1151 GGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTAC 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 GGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTAC 248

Qy       1211 TGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGT 1270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 TGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGT 308

Qy       1271 AAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTC 1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 AAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTC 368

Qy       1331 AGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGC 1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 AGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGC 428

Qy       1391 GAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCG 1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 GAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCG 488

Qy       1451 TGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGAC 1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 TGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGAC 548

Qy       1511 CCGCTGGGCCTTTGGTCTGCATCCC 1535
              |||||||||||||||||||||||||
Db        549 CCGCTGGGCCTTTGGTCTGCATCCC 573


RESULT 13
AW271767/c
LOCUS       AW271767                 547 bp    mRNA    linear   EST 03-JAN-2000
DEFINITION  xs18a11.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2769980 3',
            mRNA sequence.
ACCESSION   AW271767
VERSION     AW271767.1  GI:6658797
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 443.
FEATURES             Location/Qualifiers
     source          1..547
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2769980"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid11"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
ORIGIN

  Query Match             19.8%;  Score 546;  DB 2;  Length 547;
  Best Local Similarity   99.8%;  Pred. No. 2.5e-132;
  Matches  546;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2179 GGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGG 2238
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GGTTGAGCTTCNTGGTGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGG 488

Qy       2239 TGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGT 2298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 TGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGT 428

Qy       2299 GTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGAC 2358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGAC 368

Qy       2359 TGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCT 2418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 TGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCT 308

Qy       2419 CTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCC 2478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCC 248

Qy       2479 AGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTG 2538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 AGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTG 188

Qy       2539 GATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTTGTCCT 2598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 GATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTTGTCCT 128

Qy       2599 ACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTG 2658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 ACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTG 68

Qy       2659 CGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAATGCAAATAAAACTCACAT 2718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 CGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAATGCAAATAAAACTCACAT 8

Qy       2719 TTGTAAG 2725
              |||||||
Db          7 TTGTAAG 1
```

```
RESULT 14
AW173671/c
LOCUS       AW173671                 535 bp    mRNA    linear   EST 16-NOV-1999
DEFINITION  xj10g08.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2656862 3',
            mRNA sequence.
ACCESSION   AW173671
VERSION     AW173671.1  GI:6439619
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 535)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 420.
FEATURES             Location/Qualifiers
     source          1..535
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2656862"
                     /tissue_type="moderately-differentiated endometrial
                     adenocarcinoma, 3 pooled tumors"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Ut2"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.85 kb. Life Technologies catalog #:
                     11539-012"
ORIGIN

  Query Match             19.3%;  Score 532.4;  DB 2;  Length 535;
  Best Local Similarity   99.6%;  Pred. No. 9.6e-129;
  Matches  533;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2191 TGGTGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGACAAGTTC 2250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 TGGTGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGACAAGTTC 476

Qy       2251 TTTTGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGG 2310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 TTTTGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGG 416

Qy       2311 GAAAAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGC 2370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 GAAAAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGC 356

Qy       2371 CTCTCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATG 2430
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 CTCTCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATG 296

Qy       2431 TATTTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGAC 2490
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 TATTTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGAC 236

Qy       2491 TCTTTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGT 2550
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 TCTTTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGT 176

Qy       2551 GGGTGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTT 2610
              |||||| | |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 GGGTGTAANAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTT 116

Qy       2611 TTGTACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTGCGGACGAGAAGT 2670
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 TTGTACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTGCGGACGAGAAGT 56

Qy       2671 CTGTATGTGGGATCTGTGCTTGGGTTAGAATGCAAATAAAACTCACATTTGTAAG 2725
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         55 CTGTATGTGGGATCTGTGCTTGGGTTAGAATGCAAATAAAACTCACATTTGTAAG 1


RESULT 15
AW195353/c
LOCUS       AW195353                 538 bp    mRNA    linear   EST 29-NOV-1999
DEFINITION  xn37d01.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2695873 3',
            mRNA sequence.
ACCESSION   AW195353
VERSION     AW195353.1  GI:6474429
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 464.
FEATURES             Location/Qualifiers
     source          1..538
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2695873"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid11"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "
ORIGIN

  Query Match             19.1%;  Score 526;  DB 2;  Length 538;
  Best Local Similarity   99.8%;  Pred. No. 4.7e-127;
  Matches  537;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        475 TGGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTG 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 TGGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTG 479

Qy        535 TGCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAG 594
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          478 TGCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAG 419

Qy          595 GCAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTG 654
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          418 GCAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTG 359

Qy          655 ATGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATG 714
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          358 ATGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATG 299

Qy          715 GGGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTT 774
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          298 GGGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTT 239

Qy          775 CGTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCA 834
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          238 CGTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCA 179

Qy          835 AAGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTG 894
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          178 AAGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTG 119

Qy          895 ACTGCCCTTGTCAGCTAG-CATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCAC 953
                |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db          118 ACTGCCCTTGTCAGCTAGNCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCAC 59

Qy          954 TGATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCAC 1011
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           58 TGATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCAC 1

Search completed: November 29, 2004, 14:15:24
Job time : 8397 secs
```

```
                     GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         November 29, 2004, 08:42:07 ; Search time 2334 Seconds
                                          (without alignments)
                                          6394.689 Million cell updates/sec

Title:          US-09-989-920-100
Perfect score:  2754
Sequence:       1 gccagaagcagcctcagctt..........aaaaataaaaagatgcggcc 2754

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       3640242 seqs, 2709731945 residues

Total number of hits satisfying chosen parameters:      7280484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
                1:  /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
                14:  /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
                16:  /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
                17:  /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
                18:  /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                19:  /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
                20:  /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
                21:  /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result           Query
  No.    Score   Match Length DB  ID                         Description
 -------------------------------------------------------------------------
     1    2754   100.0   2754  9  US-09-989-920-100          Sequence 100, App
     2    2754   100.0   2754 14  US-10-074-475-2            Sequence 2, Appli
 c   3  2397.4    87.1   2399  9  US-09-989-920-91           Sequence 91, Appl
     4   612.4    22.2    960  9  US-09-989-920-99           Sequence 99, Appl
 c   5     524    19.0    524 15  US-10-172-118-2364         Sequence 2364, Ap
 c   6     524    19.0    524 16  US-10-342-887-2364         Sequence 2364, Ap
     7   483.8    17.6    591 14  US-10-074-475-1            Sequence 1, Appli
     8   357.4    13.0    395 16  US-10-609-021-176          Sequence 176, App
     9     295    10.7    295 15  US-10-116-712-429          Sequence 429, App
 c  10   224.4     8.1    251  9  US-09-989-920-90           Sequence 90, Appl
 c  11     161     5.8    161 15  US-10-102-524-1276         Sequence 1276, Ap
    12    93.6     3.4   2285 10  US-09-854-867-10           Sequence 10, Appl
 c  13      75     2.7  32038  9  US-09-764-878-292          Sequence 292, App
 c  14      75     2.7  32038 14  US-10-079-854-292          Sequence 292, App
    15      75     2.7  32844 11  US-09-984-429-498          Sequence 498, App
 c  16      75     2.7  72678 17  US-10-322-281-214          Sequence 214, App
    17    70.8     2.6  10410 17  US-10-302-027-4            Sequence 4, Appli
 c  18      66     2.4 215221 13  US-10-087-192-1360         Sequence 1360, Ap
 c  19    65.2     2.4 248436 13  US-10-087-192-2014         Sequence 2014, Ap
    20    62.8     2.3    494 13  US-10-027-632-320960       Sequence 320960,
    21    62.8     2.3    494 15  US-10-027-632-320960       Sequence 320960,
    22    62.8     2.3    498 13  US-10-027-632-96975        Sequence 96975, A
    23    62.8     2.3    498 15  US-10-027-632-96975        Sequence 96975, A
    24      61     2.2   2194 13  US-10-027-632-97802        Sequence 97802, A
    25      61     2.2   2194 15  US-10-027-632-97802        Sequence 97802, A
 c  26    56.2     2.0  10472 11  US-09-984-429-501          Sequence 501, App
 c  27    56.2     2.0 240825 10  US-09-790-289-1            Sequence 1, Appli
 c  28    56.2     2.0 240825 17  US-10-468-582-1            Sequence 1, Appli
    29    55.6     2.0 248436 13  US-10-087-192-2014         Sequence 2014, Ap
 c  30    52.2     1.9    518 13  US-10-027-632-60622        Sequence 60622, A
 c  31    52.2     1.9    518 15  US-10-027-632-60622        Sequence 60622, A
    32    52.2     1.9    665 13  US-10-027-632-42447        Sequence 42447, A
    33    52.2     1.9    665 15  US-10-027-632-42447        Sequence 42447, A
 c  34      52     1.9    459 13  US-10-027-632-87829        Sequence 87829, A
 c  35      52     1.9    459 13  US-10-027-632-87830        Sequence 87830, A
 c  36      52     1.9    459 15  US-10-027-632-87829        Sequence 87829, A
 c  37      52     1.9    459 15  US-10-027-632-87830        Sequence 87830, A
 c  38    51.4     1.9  75033 17  US-10-319-915-18           GENERAL INFORMATI
 c  39    51.2     1.9  53222 18  US-10-331-053-76           Sequence 76, Appl
 c  40    50.8     1.8   1940 15  US-10-094-749-167          Sequence 167, App
 c  41    50.8     1.8   2337 16  US-10-467-433-25           Sequence 25, Appl
    42    50.6     1.8   9544 10  US-09-764-891-7487         Sequence 7487, Ap
    43      50     1.8    657 13  US-10-027-632-136201       Sequence 136201,
    44      50     1.8    657 15  US-10-027-632-136201       Sequence 136201,
 c  45      50     1.8    935 13  US-10-027-632-120197       Sequence 120197,


                              ALIGNMENTS


RESULT 1
US-09-989-920-100
; Sequence 100, Application US/09989920
; Patent No. US20020172957A1
; GENERAL INFORMATION:
;  APPLICANT: Macina, Roberto
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Chen, Sei-Yu
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
;  FILE REFERENCE: DEX-0291
;  CURRENT APPLICATION NUMBER: US/09/989,920
;  CURRENT FILING DATE:  2001-11-21
;  PRIOR APPLICATION NUMBER: 60/252,500
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 284
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 100
;   LENGTH: 2754
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-989-920-100

  Query Match             100.0%;  Score 2754;  DB 9;  Length 2754;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2754;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60

Qy         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120
```

```
Qy    121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180

Qy    181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240

Qy    241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300

Qy    301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360

Qy    361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420

Qy    421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480

Qy    481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540

Qy    541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600

Qy    601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660

Qy    661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720

Qy    721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 780

Qy    781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840

Qy    841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900

Qy    901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960

Qy    961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020

Qy   1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080

Qy   1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140

Qy   1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200

Qy   1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260

Qy   1261 CTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320

Qy   1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380

Qy   1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440

Qy   1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500

Qy   1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560

Qy   1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620

Qy   1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680

Qy   1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740

Qy   1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800

Qy   1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860

Qy   1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920

Qy   1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980

Qy   1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040

Qy   2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100

Qy   2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160

Qy   2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220

Qy   2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280

Qy   2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340
```

```
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340

Qy          2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400

Qy          2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460

Qy          2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGAGTATCTGTC 2520
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGAGTATCTGTC 2520

Qy          2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580

Qy          2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640

Qy          2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700

Qy          2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754


RESULT 2
US-10-074-475-2
; Sequence 2, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
;  APPLICANT: Salceda, Susana
;  APPLICANT:  Macina, Roberto
;  APPLICANT:  Hu, Ping
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Karra, Kalpana
;  APPLICANT:  Cafferkey, Robert
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
;  TITLE OF INVENTION:  Genes and Proteins
;  FILE REFERENCE: DEX-0313
;  CURRENT APPLICATION NUMBER: US/10/074,475
;  CURRENT FILING DATE:  2002-02-13
;  PRIOR APPLICATION NUMBER: 60/268,292
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 295
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 2754
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-074-475-2

  Query Match            100.0%;  Score 2754;  DB 14;  Length 2754;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2754;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy             1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60

Qy            61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120

Qy           121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180

Qy           181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240

Qy           241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300

Qy           301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360

Qy           361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420

Qy           421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480

Qy           481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540

Qy           541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600

Qy           601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660

Qy           661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720

Qy           721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 780

Qy           781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840

Qy           841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900

Qy           901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960

Qy           961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020

Qy          1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080

Qy          1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140

Qy          1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200

Qy          1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260

Qy        1261 CTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 CTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320

Qy        1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380

Qy        1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440

Qy        1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500

Qy        1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560

Qy        1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620

Qy        1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680

Qy        1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740

Qy        1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800

Qy        1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860

Qy        1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920

Qy        1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980

Qy        1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040

Qy        2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100

Qy        2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160

Qy        2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220

Qy        2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280

Qy        2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340

Qy        2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400

Qy        2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460

Qy        2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520

Qy        2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580

Qy        2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640

Qy        2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700

Qy        2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754


RESULT 3
US-09-989-920-91/c
; Sequence 91, Application US/09989920
; Patent No. US20020172957A1
; GENERAL INFORMATION:
;  APPLICANT: Macina, Roberto
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Chen, Sei-Yu
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
;  FILE REFERENCE: DEX-0291
;  CURRENT APPLICATION NUMBER: US/09/989,920
;  CURRENT FILING DATE:  2001-11-21
;  PRIOR APPLICATION NUMBER: 60/252,500
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 284
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 91
;   LENGTH: 2399
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-989-920-91

  Query Match            87.1%;  Score 2397.4;  DB 9;  Length 2399;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 2398;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         356 TTGAGGGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGC 415
               |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        2399 TTGAGGGCCAGTGAGGTGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGC 2340

Qy         416 ACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCT 475
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2339 ACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCT 2280

Qy         476 GGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGT 535
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2279 GGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGT 2220

Qy         536 GCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGG 595
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2219 GCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGG 2160
```

```
Qy        596 CAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGA 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2159 CAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGA 2100

Qy        656 TGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2099 TGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGG 2040

Qy        716 GGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTC 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2039 GGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTC 1980

Qy        776 GTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAA 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1979 GTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAA 1920

Qy        836 AGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGA 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1919 AGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGA 1860

Qy        896 CTGCCCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTG 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1859 CTGCCCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTG 1800

Qy        956 ATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAA 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1799 ATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAA 1740

Qy       1016 AAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTA 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1739 AAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTA 1680

Qy       1076 TTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1679 TTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATA 1620

Qy       1136 TTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAA 1195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1619 TTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAA 1560

Qy       1196 AAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAG 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1559 AAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAG 1500

Qy       1256 AAAGCCTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAA 1315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 AAAGCCTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAA 1440

Qy       1316 GCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTG 1375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1439 GCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTG 1380

Qy       1376 TCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGA 1435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 TCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGA 1320

Qy       1436 GTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTT 1495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 GTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTT 1260

Qy       1496 CCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGT 1555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 CCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGT 1200

Qy       1556 TGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCAC 1615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 TGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCAC 1140

Qy       1616 ACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGG 1675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 ACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGG 1080

Qy       1676 ACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGC 1735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 ACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGC 1020

Qy       1736 TACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGC 1795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 TACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGC 960

Qy       1796 CCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCC 1855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCC 900

Qy       1856 TCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCC 1915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 TCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCC 840

Qy       1916 CTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGC 1975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 CTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGC 780

Qy       1976 CACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAG 2035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 CACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAG 720

Qy       2036 GCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAG 2095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 GCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAG 660

Qy       2096 GCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCC 2155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 GCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCC 600

Qy       2156 ACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTG 2215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 ACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTG 540

Qy       2216 CAGATCTCATGCTCTCAGATAGGTGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCC 2275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CAGATCTCATGCTCTCAGATAGGTGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCC 480

Qy       2276 TGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGG 2335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGG 420

Qy       2336 TTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGG 2395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGG 360

Qy       2396 TGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAG 2455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAG 300

Qy       2456 ATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTAT 2515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTAT 240

Qy       2516 CTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTG 2575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTG 180

Qy       2576 GGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTT 2635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 GGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTT 120

Qy       2636 GCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGT 2695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 GCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGT 60

Qy       2696 TAGAATGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 TAGAATGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAATAAAAAGATGCGGCC 1
```

```
RESULT 4
US-09-989-920-99
; Sequence 99, Application US/09989920
; Patent No. US20020172957A1
; GENERAL INFORMATION:
;  APPLICANT: Macina, Roberto
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Chen, Sei-Yu
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
;  FILE REFERENCE: DEX-0291
;  CURRENT APPLICATION NUMBER: US/09/989,920
;  CURRENT FILING DATE:  2001-11-21
;  PRIOR APPLICATION NUMBER: 60/252,500
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 284
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
;   LENGTH: 960
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (716)..(716)
;   OTHER INFORMATION: a, c, g or t
US-09-989-920-99

  Query Match             22.2%;  Score 612.4;  DB 9;  Length 960;
  Best Local Similarity   99.4%;  Pred. No. 8.4e-157;
  Matches  646;  Conservative    0;  Mismatches    1;  Indels    3;  Gaps    3;

Qy        1658 GTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTG 1717
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 GTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTG 100

Qy        1718 TCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTT 1777
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         101 TCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTT 160

Qy        1778 GAGGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCC 1837
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         161 GAGGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCC 220

Qy        1838 ATTGGAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGG 1897
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         221 ATTGGAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGG 280

Qy        1898 CTGTGGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACT 1957
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         281 CTGTGGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACT 340

Qy        1958 CCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAG 2017
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         341 CCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAG 400

Qy        2018 CTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCT 2077
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         401 CTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCT 460

Qy        2078 CCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTA 2137
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         461 CCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTA 520

Qy        2138 GATGGGGGCCAGGGTCCCACA-GCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGT 2196
               |||||||||||||||||||| ||| |||||||||||||||||||||||||||||||||||
Db         521 GATGGGGGCCAGGGTCCCACATGCTGCTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGT 580

Qy        2197 CCAGTCCGGAT-CCCACTTGCAGATCTCATGCTCTCAGATAGGTGGACAAGTTCTTTTG 2255
               ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         581 CCAGTCCGGATCCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGACAAGTTCTTTTG 640

Qy        2256 TCACAGTGCTGG-CTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTC 2304
               |||||||||||| |||||||||||||||||||||||||||||||||||||
Db         641 TCACAGTGCTGGCCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTC 690


RESULT 5
US-10-172-118-2364/c
; Sequence 2364, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Chris
;  APPLICANT:  Van 't Veer, Laura
;  APPLICANT:  Van de Vijver, Marc
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-175-999
;  CURRENT APPLICATION NUMBER: US/10/172,118
;  CURRENT FILING DATE:  2002-06-14
;  PRIOR APPLICATION NUMBER: 60/380,770
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 2364
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;  PUBLICATION INFORMATION:
;   DATABASE ACCESSION NUMBER: Contig46362
;   DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-2364

  Query Match             19.0%;  Score 524;  DB 15;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 1e-132;
  Matches  524;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         488 TCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAGTGGCCAG 547
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         524 TCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAGTGGCCAG 465

Qy         548 GGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTGTCCTTAG 607
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         464 GGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTGTCCTTAG 405

Qy         608 GAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGATAAGGTG 667
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         404 GAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGATAAGGTG 345

Qy         668 GGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAACCAAACC 727
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         344 GGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAACCAAACC 285

Qy         728 ATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCTGAGGAGC 787
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         284 ATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCTGAGGAGC 225

Qy         788 CCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGAACAACAT 847
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         224 CCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGAACAACAT 165

Qy         848 CCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTCA 907
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         164 CCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTCA 105

Qy         908 GCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTATGATAGG 967
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         104 GCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTATGATAGG 45

Qy         968 GAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCAC 1011
               ||||||||||||||||||||||||||||||||||||||||||||
Db          44 GAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCAC 1
```

```
RESULT 6
US-10-342-887-2364/c
; Sequence 2364, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
;  APPLICANT: Dai, Hongyue
;  APPLICANT:  He, Yudong
;  APPLICANT:  Linsley, Peter S.
;  APPLICANT:  Mao, Mao
;  APPLICANT:  Roberts, Christopher J.
;  APPLICANT:  Van 't Veer, Laura Johanna
;  APPLICANT:  Van de Vijver, Marc J.
;  APPLICANT:  Bernards, Rene
;  TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
;  FILE REFERENCE: 9301-188-999
;  CURRENT APPLICATION NUMBER: US/10/342,887
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: 60/298,918
;  PRIOR FILING DATE: 2001-06-18
;  PRIOR APPLICATION NUMBER: 60/380,710
;  PRIOR FILING DATE: 2002-05-14
;  PRIOR APPLICATION NUMBER: 10/172,118
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 2364
;   LENGTH: 524
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-342-887-2364

  Query Match             19.0%;  Score 524;  DB 16;  Length 524;
  Best Local Similarity   100.0%;  Pred. No. 1e-132;
  Matches  524;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        488 TCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAGTGGCCAG 547
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 TCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAGTGGCCAG 465

Qy        548 GGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTGTCCTTAG 607
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 GGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTGTCCTTAG 405

Qy        608 GAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGATAAGGTG 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 GAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGATAAGGTG 345

Qy        668 GGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAACCAAACC 727
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 GGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAACCAAACC 285

Qy        728 ATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCTGAGGAGC 787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 ATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCTGAGGAGC 225

Qy        788 CCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGAACAACAT 847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGAACAACAT 165

Qy        848 CCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTCA 907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 CCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTCA 105

Qy        908 GCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTATGATAGG 967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 GCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTATGATAGG 45

Qy        968 GAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCAC 1011
              ||||||||||||||||||||||||||||||||||||||||||||
Db         44 GAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCAC 1

RESULT 7
```

```
US-10-074-475-1
; Sequence 1, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
;  APPLICANT: Salceda, Susana
;  APPLICANT:  Macina, Roberto
;  APPLICANT:  Hu, Ping
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Karra, Kalpana
;  APPLICANT:  Cafferkey, Robert
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
;  TITLE OF INVENTION:  Genes and Proteins
;  FILE REFERENCE: DEX-0313
;  CURRENT APPLICATION NUMBER: US/10/074,475
;  CURRENT FILING DATE:  2002-02-13
;  PRIOR APPLICATION NUMBER: 60/268,292
;  PRIOR FILING DATE: 2001-02-13
;  NUMBER OF SEQ ID NOS: 295
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 591
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-074-475-1

  Query Match             17.6%;  Score 483.8;  DB 14;  Length 591;
  Best Local Similarity   97.5%;  Pred. No. 1.2e-121;
  Matches  555;  Conservative    0;  Mismatches    7;  Indels    7;  Gaps    6;

Qy       1735 CTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGG 1794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 CTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGG 71

Qy       1795 CCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGA-GGATGGACAG 1853
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||
Db         72 CCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGATGGATGTACAG 131

Qy       1854 CCTCAAATGGAAG-GAGTCCCACGGGAGATGGGTCCGAGGTCC-GGCTGTGGCCATCCAG 1911
              ||||||||||||| ||||||||||||||||||||||||||||| ||||||||||||||||
Db        132 CCTCAAATGGAAGTGAGTCCCACGGGAGATGGGTCCGAGGTCCAGGCTGTGGCCATCCAG 191

Qy       1912 CCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAG 1971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAG 251

Qy       1972 CAGCCACTGCAGTTCCTTTCTTCGTGAG-TAACAGTAGTGATAGCAGCTGGGGCTAACAG 2030
               ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        252 TAGCCACTGCAGTTCCTTTCTTCGTGAGATAACAGTAGTGATAGCAGCTGGGGCTAACAG 311

Qy       2031 GCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATG 2090
              ||||||||| |||| |||||||||||||||||||||||||||||||||||||||||||||
Db        312 GCTAGGCTTAGTGTCCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATG 371

Qy       2091 GGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAG- 2149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        372 GGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGA 431

Qy       2150 GGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTT--CCTGGTGTCCAGTCCGGAT 2207
              |||||||||||||||||||||||||||||||||||||||  |  | ||||||||||||||
Db        432 GGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTTCCTGGATGTCCAGTCCGGAT 491

Qy       2208 CCCACTTGCAGATCTCATGCTCTCAGATAGGTGGACAAGTTCTTTTGTCACAGTGCTGG 2267
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 CCCACTTGCAGATCTCATGCTCTCAGATAGGTGGACAAGTTCTTTTGTCACAGTGCTGG 551

Qy       2268 CTCTGTCCTGAGGCCTCATTGCTGGCTGG 2296
              |||||||||||||||||||||||||||||
Db        552 CTCTGTCCTGAGGCCTCATTGCTGGCTGG 580

RESULT 8
```

```
US-10-609-021-176
; Sequence 176, Application US/10609021
; Publication No. US20040086913A1
; GENERAL INFORMATION:
;  APPLICANT: Williams, Lewis T.
;  APPLICANT:  Escobedo, Jaime
;  APPLICANT:  Innis, Michael A.
;  APPLICANT:  Garcia, Pablo Dominiguez
;  APPLICANT:  Sudduth-Klinger, Julie
;  APPLICANT:  Reinhard, Christoph
;  APPLICANT:  He, Zhijun
;  APPLICANT:  Randazzo, Filippo
;  APPLICANT:  Kennedy, Giulia C.
;  APPLICANT:  Pot, David
;  APPLICANT:  Kassam, Altaf
;  APPLICANT:  Lamson, George
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Jones, Lee William
;  TITLE OF INVENTION: Human Genes and Gene Expression Products
;  TITLE OF INVENTION:  XVI
;  FILE REFERENCE: 2300-1625CON
;  CURRENT APPLICATION NUMBER: US/10/609,021
;  CURRENT FILING DATE:  2003-06-26
;  PRIOR APPLICATION NUMBER: 60/192,583
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 09/819,150
;  PRIOR FILING DATE: 2001-03-27
;  NUMBER OF SEQ ID NOS: 324
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 176
;   LENGTH: 395
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-609-021-176

  Query Match             13.0%;  Score 357.4;  DB 16;  Length 395;
  Best Local Similarity   96.9%;  Pred. No. 4.2e-87;
  Matches  375;  Conservative    0;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       1216 CAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGG 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 CAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGG 68

Qy       1276 C-TTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGAT 1334
              |  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CTTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGAT 128

Qy       1335 GAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGC 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 GAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGC 188

Qy       1395 CCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGT 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 CCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGT 248

Qy       1455 GGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGC 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 GGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGC 308

Qy       1515 TGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACT 1574
              |||||||||||||||||||||||||||||||||||||||| ||||||| | |||||||||
Db        309 TGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGCTGATGCGCGTAGAAGGACT 368

Qy       1575 TTGAGCAGTGGTGGGCAGCAGTGGCCT 1601
              ||||||||||||||       || |||
Db        369 TTGAGCAGTGGTGGCAGCACTTGCCCT 395


RESULT 9
US-10-116-712-429
; Sequence 429, Application US/10116712
; Publication No. US20030194764A1
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Switzer, Ann
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.568
;  CURRENT APPLICATION NUMBER: US/10/116,712
;  CURRENT FILING DATE:  2002-04-07
;  NUMBER OF SEQ ID NOS: 670
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 429
;   LENGTH: 295
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-116-712-429

  Query Match             10.7%;  Score 295;  DB 15;  Length 295;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-70;
  Matches  295;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1723 CTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGG 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGG 60

Qy       1783 GCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGG 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGG 120

Qy       1843 AGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTG 1902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTG 180

Qy       1903 GCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGG 1962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGG 240

Qy       1963 GGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAG 2017
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAG 295


RESULT 10
US-09-989-920-90/c
; Sequence 90, Application US/09989920
; Patent No. US20020172957A1
; GENERAL INFORMATION:
;  APPLICANT: Macina, Roberto
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Chen, Sei-Yu
;  APPLICANT:  Sun, Yongming
;  APPLICANT:  Liu, Chenghua
;  TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
;  FILE REFERENCE: DEX-0291
;  CURRENT APPLICATION NUMBER: US/09/989,920
;  CURRENT FILING DATE:  2001-11-21
;  PRIOR APPLICATION NUMBER: 60/252,500
;  PRIOR FILING DATE: 2000-11-22
;  NUMBER OF SEQ ID NOS: 284
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 90
;   LENGTH: 251
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-989-920-90

  Query Match             8.1%;  Score 224.4;  DB 9;  Length 251;
  Best Local Similarity   98.8%;  Pred. No. 9.6e-51;
  Matches  237;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy       2381 CCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAAT 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 CCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAAT 192
```

```
Qy       2441 GATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAA 2500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAA 132

Qy       2501 AGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGA 2560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 AGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGA 72

Qy       2561 GGTCCTGGGGAGGT--GGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCT 2618
              ||||||||||||||  | ||||||||||||||||||||||||||||||||||||||||||
Db         71 GGTCCTGGGGAGGTGCGCGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCT 12


RESULT 11
US-10-102-524-1276/c
; Sequence 1276, Application US/10102524
; Publication No. US20030109434A1
; GENERAL INFORMATION:
;  APPLICANT: Algate, Paul A.
;  APPLICANT:  Mannion, Jane
;  APPLICANT:  Gaiger, Alexander
;  APPLICANT:  Gordon, Brian
;  APPLICANT:  Harlocker, Susan L.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION:  THERAPY AND DIAGNOSIS OF KIDNEY CANCER
;  FILE REFERENCE: 210121.572
;  CURRENT APPLICATION NUMBER: US/10/102,524
;  CURRENT FILING DATE:  2002-03-19
;  NUMBER OF SEQ ID NOS: 1863
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1276
;   LENGTH: 161
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-102-524-1276

  Query Match             5.8%;  Score 161;  DB 15;  Length 161;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-33;
  Matches  161;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1727 TCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCG 1786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 TCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCG 102

Qy       1787 CCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGA 1846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 CCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGA 42

Qy       1847 TGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTC 1887
              |||||||||||||||||||||||||||||||||||||||||
Db         41 TGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTC 1


RESULT 12
US-09-854-867-10
; Sequence 10, Application US/09854867
; Publication No. US20030224356A1
; GENERAL INFORMATION:
;  APPLICANT: JOAN, KNOLL H
;  APPLICANT:  ROGAN, PETER K
;  TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING
;  FILE REFERENCE: 30307
;  CURRENT APPLICATION NUMBER: US/09/854,867
;  CURRENT FILING DATE:  2003-05-08
;  NUMBER OF SEQ ID NOS: 613
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 2285
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: repeat_region
;   LOCATION: (1)..(2285)
;   OTHER INFORMATION: cheshire
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (377)..(377)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (580)..(580)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (651)..(651)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (664)..(664)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (725)..(725)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (728)..(728)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1095)..(1095)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1151)..(1151)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1281)..(1281)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1351)..(1351)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1370)..(1370)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1385)..(1385)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1885)..(1885)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1912)..(1912)
;   OTHER INFORMATION: n is a, c, g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1990)..(1990)
;   OTHER INFORMATION: n is a, c, g or t
US-09-854-867-10

  Query Match             3.4%;  Score 93.6;  DB 10;  Length 2285;
  Best Local Similarity   67.3%;  Pred. No. 2.5e-14;
  Matches  183;  Conservative    1;  Mismatches   70;  Indels   18;  Gaps    3;

Qy        957 TTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAA 1016
              || ||||||||||| | || |||||||||| |   |             |   | ||| |
Db       2016 TTGATGATAGGGAACACTAACTTTGAACCCCAATTAAGCAAAATGTTATCCCWAAAAAGA 2075

Qy       1017 AAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTAT 1076
```

```
              |    |||||| |  ||||||||| ||||||   ||||||  |||||||||||||| | :
Db       2076 ATTCCATTCTTCTCATTAGTAGACCTGTATTA--CAAAAAATTGTACTCAATTATTATTW 2133

Qy       1077 TTTGGATTTTATCAATT-------TAAAAATTGTGGAAATTTGTTTGCTCTTA------- 1122
              ||   || || | ||||       |||||||||||||| ||| | |  | |||
Db       2134 TTATTATATTTTGAATTTCATCAATAAAAATTGTGGAATTTTTTCTCTTGTTATATAAGT 2193

Qy       1123 --CGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCT 1180
                |  ||  |||   |||||||||||||||||| |  |||||| |||||||||||  |||
Db       2194 ACCTACATAATATCCTTGATTTTGCCTCTTGGCCCGCAAAGCCTAAAATATTTACTATCT 2253

Qy       1181 AGCCCGTTACAGAAAAAGTCTGCTGACTACTG 1212
               |||| ||||||||||||| ||| |||  |||
Db       2254 GGCCCTTTACAGAAAAAGTTTGCCGACCCCTG 2285


RESULT 13
US-09-764-878-292/c
; Sequence 292, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121
;  CURRENT APPLICATION NUMBER: US/09/764,878
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 292
;   LENGTH: 32038
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-878-292

  Query Match             2.7%;  Score 75;  DB 9;  Length 32038;
  Best Local Similarity   65.9%;  Pred. No. 1.5e-08;
  Matches  147;  Conservative    0;  Mismatches   60;  Indels   16;  Gaps    2;

Qy       1013 AAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATT 1072
              ||| |   ||||||| |  |||||| ||  |||||    |||||  ||| || |||||||
Db      28877 AAATATTTAATTCTTTTCATTAGTAAACCCGTATTACAAAAAAAATTGTTCTAAATTATT 28818

Qy       1073 GTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGC------- 1125
              |  ||||| ||||| |||||  ||   | ||| ||||||  || ||||   |
Db      28817 G--TTTTGAATTTTGTCAATAAAATTTTCGTGCAAATTTTATTTCTCTCTAGTTATTTAA 28760

Qy       1126 -------CAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGT 1178
                     ||  | ||| |   |||| |||||| || |  |||||| |||||||||||  |
Db      28759 GTATCTACATAACAATCTCAGTTTTACCTCTTAGCCCACAAAGCCTAAAATATTTACTAT 28700

Qy       1179 CTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
              || |||| ||| ||||||||| | |||||  |||  | ||| |
Db      28699 CTGGCCCTTTATAGAAAAAGTTTCCTGACCTCTGCCCTAGAGC 28657


RESULT 14
US-10-079-854-292/c
; Sequence 292, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121C1
;  CURRENT APPLICATION NUMBER: US/10/079,854
;  CURRENT FILING DATE:  2002-02-22
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 292
;   LENGTH: 32038
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
US-10-079-854-292

  Query Match             2.7%;  Score 75;  DB 14;  Length 32038;
  Best Local Similarity   65.9%;  Pred. No. 1.5e-08;
  Matches  147;  Conservative    0;  Mismatches   60;  Indels   16;  Gaps    2;

Qy       1013 AAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATT 1072
              ||| |   ||||||| |  |||||| ||  |||||    |||||  ||| || |||||||
Db      28877 AAATATTTAATTCTTTTCATTAGTAAACCCGTATTACAAAAAAAATTGTTCTAAATTATT 28818

Qy       1073 GTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGC------- 1125
              |  ||||| ||||| |||||  ||   | ||| ||||||  || ||||   |
Db      28817 G--TTTTGAATTTTGTCAATAAAATTTTCGTGCAAATTTTATTTCTCTCTAGTTATTTAA 28760

Qy       1126 -------CAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGT 1178
                     ||  | ||| |   |||| |||||| || |  |||||| |||||||||||  |
Db      28759 GTATCTACATAACAATCTCAGTTTTACCTCTTAGCCCACAAAGCCTAAAATATTTACTAT 28700

Qy       1179 CTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
              || |||| ||| ||||||||| | |||||  |||  | ||| |
Db      28699 CTGGCCCTTTATAGAAAAAGTTTCCTGACCTCTGCCCTAGAGC 28657


RESULT 15
US-09-984-429-498
; Sequence 498, Application US/09984429
; Publication No. US20040010132A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 53 Human Secreted Proteins
;  FILE REFERENCE: PZ018P2
;  CURRENT APPLICATION NUMBER: US/09/984,429
;  CURRENT FILING DATE:  2001-10-30
;  PRIOR APPLICATION NUMBER: 60/244,591
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 09/288,143
;  PRIOR FILING DATE: 1999-04-08
;  PRIOR APPLICATION NUMBER: PCT/US98/21142
;  PRIOR FILING DATE: 1998-10-08
;  PRIOR APPLICATION NUMBER: 60/061,463
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,529
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/071,498
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,527
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,536
;  PRIOR FILING DATE: 1997-10-09
;  PRIOR APPLICATION NUMBER: 60/061,532
;  PRIOR FILING DATE: 1997-10-09
;  NUMBER OF SEQ ID NOS: 727
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 498
;   LENGTH: 32844
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-984-429-498

  Query Match             2.7%;  Score 75;  DB 11;  Length 32844;
  Best Local Similarity   65.9%;  Pred. No. 1.5e-08;
  Matches  147;  Conservative    0;  Mismatches   60;  Indels   16;  Gaps    2;

Qy       1013 AAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATT 1072
              ||| |   ||||||| |  |||||| ||  |||||    |||||  ||| || |||||||
Db       3968 AAATATTTAATTCTTTTCATTAGTAAACCCGTATTACAAAAAAAATTGTTCTAAATTATT 4027

Qy       1073 GTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGC------- 1125
              |  ||||| ||||| |||||  ||   | ||| ||||||  || ||||   |
Db       4028 G--TTTTGAATTTTGTCAATAAAATTTTCGTGCAAATTTTATTTCTCTCTAGTTATTTAA 4085

Qy       1126 -------CAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGT 1178
                     ||  | ||| |   |||| |||||| || |  |||||| |||||||||||  |
```

```
Db       4086 GTATCTACATAACAATCTCAGTTTTACCTCTTAGCCCACAAAGCCTAAAATATTTACTAT 4145

Qy       1179 CTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
              || |||| ||| ||||||||| | |||||  |||  | ||| |
Db       4146 CTGGCCCTTTATAGAAAAAGTTTCCTGACCTCTGCCCTAGAGC 4188

Search completed: November 29, 2004, 15:19:46
Job time : 2343 secs
```

```
                       GenCore version 5.1.6
             Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          November 29, 2004, 08:42:07 ; Search time 1281 Seconds
                                            (without alignments)
                                           11285.641 Million cell updates/sec

Title:           US-09-989-920-100
Perfect score:   2754
Sequence:        1 gccagaagcagcctcagctt..........aaaaataaaaagatgcggcc 2754

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters:      8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_23Sep04:*
                1:  geneseqn1980s:*
                2:  geneseqn1990s:*
                3:  geneseqn2000s:*
                4:  geneseqn2001as:*
                5:  geneseqn2001bs:*
                6:  geneseqn2002as:*
                7:  geneseqn2002bs:*
                8:  geneseqn2003as:*
                9:  geneseqn2003bs:*
                10:  geneseqn2003cs:*
                11:  geneseqn2003ds:*
                12:  geneseqn2004s:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
                   %
Result             Query
  No.    Score     Match Length DB  ID                           Description
 ------------------------------------------------------------------------------
    1     2754   100.0    2754  6  ABX92058                     Abx92058 Lung spec
    2     2754   100.0    2754  6  ABT07523                     Abt07523 Human bre
c   3   2397.4    87.1    2399  6  ABX92049                     Abx92049 Lung spec
    4    612.4    22.2     960  6  ABX92057                     Abx92057 Lung spec
    5      554    20.1     611  5  ABV54530                     Abv54530 Human pro
    6    483.8    17.6     591  6  ABT07522                     Abt07522 Human bre
    7    357.4    13.0     395  4  AAS59058                     Aas59058 Human can
    8      295    10.7     295  10  ADD49697                     Add49697 Human lun
c   9    224.4     8.1     251  6  ABX92048                     Abx92048 Lung spec
   10     76.6     2.8     736  3  AAZ35057                     Aaz35057 Pluripote
c  11       75     2.7   32038  4  AAK90877                     Aak90877 Human dig
c  12       75     2.7   32038  4  AAK89980                     Aak89980 Human dig
c  13       75     2.7   32038  5  AAS30028                     Aas30028 Human lun
c  14       75     2.7   32038  10  ADB33365                     Adb33365 Human nov
   15       75     2.7   32844  12  ADJ12644                     Adj12644 DNA fragm
   16     70.8     2.6   10410  12  ADN37124                     Adn37124 DNA seque
   17     64.8     2.4    4153  4  AAK66121                     Aak66121 Human imm
   18     58.8     2.1  120670  12  ADQ59167                     Adq59167 MSI-H car
   19     56.6     2.1    2000  8  ADA71938                     Ada71938 Rice gene
c  20     56.2     2.0   10472  12  ADJ12647                     Adj12647 DNA fragm
c  21     56.2     2.0  240823  10  ADD69391                     Add69391 Human PG-
```

```
c  22     56.2     2.0  240825  4  AAF24497                     Aaf24497 Human PG-
c  23     56.2     2.0  240825  6  ABQ81802                     Abq81802 Human PG-
c  24     55.8     2.0  147419  6  ABK83574                     Abk83574 Human cDN
c  25       55     2.0  200620  12  ADO56277                     Ado56277 Human pre
c  26     51.4     1.9   75033  12  ADP66763                     Adp66763 Human end
c  27     50.8     1.8    1940  10  ADA52599                     Ada52599 Human cod
c  28     50.8     1.8    2193  8  ABX05077                     Abx05077 Human nov
c  29     50.8     1.8    2337  6  AAL49931                     Aal49931 Human mol
   30     50.6     1.8    9544  4  AAL04799                     Aal04799 Human rep
   31     50.6     1.8    9544  4  ABL97693                     Abl97693 Human tes
   32     50.6     1.8   32476  4  AAK85314                     Aak85314 Human imm
c  33     50.2     1.8    2000  8  ADA71938                     Ada71938 Rice gene
   34     49.4     1.8     495  6  ABN79060                     Abn79060 Human ORF
   35       49     1.8    3365  11  ADM02732                     Adm02732 Human cDN
   36       49     1.8    6330  4  AAA54431                     Aaa54431 11-cis re
   37       49     1.8    6330  10  ADD71027                     Add71027 Human ret
   38       49     1.8    6330  12  ADQ19466                     Adq19466 Human sof
   39       49     1.8   19806  6  ABN89476                     Abn89476 Human deh
   40     47.4     1.7   14024  4  AAS36381                     Aas36381 Human car
c  41     47.4     1.7   14024  5  AAS30157                     Aas30157 Human lun
c  42     47.4     1.7   14024  10  ADB33494                     Adb33494 Human nov
   43     47.4     1.7   14024  10  ADE47075                     Ade47075 Human car
   44     47.4     1.7   28965  10  ADC86716                     Adc86716 Human GPC
   45       47     1.7     265  3  AAC23203                     Aac23203 Human sec

                                 ALIGNMENTS

RESULT 1
ABX92058
ID   ABX92058 standard; cDNA; 2754 BP.
XX
AC   ABX92058;
XX
DT   08-MAY-2003  (first entry)
XX
DE   Lung specific nucleic acid (LSNA) #100.
XX
KW   Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;
KW   cancer monitoring; cancer staging; cancer imaging; lung cancer;
KW   non-cancerous diseases of the lung; transgenic animal; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200268633-A2.
XX
PD   06-SEP-2002.
XX
PF   21-NOV-2001; 2001WO-US043612.
XX
PR   22-NOV-2000; 2000US-0252500P.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Macina RA,  Recipon H,  Chen S,  Sun Y,  Liu C;
XX
DR   WPI; 2002-713376/77.
XX
PT   New isolated human nucleic acid molecule and polypeptide, useful for
PT   identifying, diagnosing, monitoring, staging, imaging and treating lung
PT   cancer and non-cancerous diseases of the lung.
XX
PS   Claim 1; Page 246-248; 389pp; English.
XX
CC   The invention describes an isolated human nucleic acid (I) encoding any
CC   of 120 10-1533 residue amino acid sequences (S1), given in the
CC   specification, comprising any of 164 179-12421 base pair sequences (S2),
CC   given in the specification. The methods and compositions of the present
CC   invention are useful for identifying, diagnosing, monitoring, staging,
CC   imaging and treating lung cancer and non-cancerous diseases of the lung.
CC   They are also used for identifying lung tissue, monitoring and
CC   identifying and/or designing antagonists of the polypeptide of the
```

```
CC   invention, gene therapy, production of transgenic animals and production
CC   of engineered lung tissue for treatment and research. This sequence
CC   encodes a lung specific nucleic acid
XX
SQ   Sequence 2754 BP; 557 A; 720 C; 753 G; 724 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 2754;  DB 6;  Length 2754;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2754;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60

Qy         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120

Qy        121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180

Qy        181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240

Qy        241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300

Qy        301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360

Qy        361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420

Qy        421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480

Qy        481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540

Qy        541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600

Qy        601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660

Qy        661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720

Qy        721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGAACCTGTCCTCATGCCCTTCGTCCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGAACCTGTCCTCATGCCCTTCGTCCT 780

Qy        781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840

Qy        841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900

Qy        901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960

Qy        961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020

Qy       1021 AATTCTTATTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATTCTTATTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080

Qy       1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140

Qy       1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200

Qy       1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260

Qy       1261 CTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320

Qy       1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380

Qy       1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440

Qy       1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500

Qy       1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560

Qy       1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620

Qy       1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680

Qy       1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740

Qy       1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800

Qy       1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860

Qy       1861 TGGAAGGAGTCCCACGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TGGAAGGAGTCCCACGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920

Qy       1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980

Qy       1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040
```

```
Qy       2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100

Qy       2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160

Qy       2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220

Qy       2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280

Qy       2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340

Qy       2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400

Qy       2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460

Qy       2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520

Qy       2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580

Qy       2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640

Qy       2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700

Qy       2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754


RESULT 2
ABT07523
ID   ABT07523 standard; cDNA; 2754 BP.
XX
AC   ABT07523;
XX
DT   14-NOV-2002  (first entry)
XX
DE   Human breast cancer associated coding sequence SEQ ID NO: 2.
XX
KW   Human; breast specific gene; breast specific protein; breast cancer;
KW   gene therapy; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200264611-A1.
XX
PD   22-AUG-2002.
XX
PF   12-FEB-2002; 2002WO-US004197.
XX
PR   13-FEB-2001; 2001US-0268292P.
XX
PA   (DIAD-) DIADEXUS INC.
```

```
XX
PI   Salceda S,  Macina RA,  Hu P,  Recipon H,  Karra K,  Cafferkey R;
PI   Sun Y,  Liu C;
XX
DR   WPI; 2002-657582/70.
XX
PT   New breast specific nucleic acids and proteins, useful for identifying,
PT   diagnosing, monitoring, staging, imaging, and treating breast cancer and
PT   non-cancerous disease states in breast tissue, and in gene therapy.
XX
PS   Claim 1; Page 148-150; 367pp; English.
XX
CC   The present invention provides human breast specific coding sequences and
CC   proteins. These can be used in the diagnosis and treatment of breast
CC   cancer and non-cancerous diseases of the breast. The present sequence is
CC   a coding sequence of the invention
XX
SQ   Sequence 2754 BP; 557 A; 720 C; 753 G; 724 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 2754;  DB 6;  Length 2754;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2754;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60

Qy         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120

Qy        121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180

Qy        181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240

Qy        241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300

Qy        301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360

Qy        361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420

Qy        421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480

Qy        481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540

Qy        541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600

Qy        601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660

Qy        661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720

Qy        721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 780

Qy          781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840

Qy          841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900

Qy          901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960

Qy          961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020

Qy         1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080

Qy         1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140

Qy         1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200

Qy         1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260

Qy         1261 CTCAGACAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1261 CTCAGACAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320

Qy         1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380

Qy         1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440

Qy         1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500

Qy         1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560

Qy         1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620

Qy         1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680

Qy         1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740

Qy         1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800

Qy         1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860

Qy         1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920

Qy         1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980

Qy         1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040

Qy         2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100

Qy         2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160

Qy         2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220

Qy         2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280

Qy         2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340

Qy         2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400

Qy         2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460

Qy         2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520

Qy         2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580

Qy         2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640

Qy         2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700

Qy         2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754


RESULT 3
ABX92049/c
ID   ABX92049 standard; cDNA; 2399 BP.
XX
AC   ABX92049;
XX
DT   08-MAY-2003  (first entry)
XX
DE   Lung specific nucleic acid (LSNA) #91.
XX
KW   Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;
```

```
KW   cancer monitoring; cancer staging; cancer imaging; lung cancer;
KW   non-cancerous diseases of the lung; transgenic animal; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200268633-A2.
XX
PD   06-SEP-2002.
XX
PF   21-NOV-2001; 2001WO-US043612.
XX
PR   22-NOV-2000; 2000US-0252500P.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Macina RA,  Recipon H,  Chen S,  Sun Y,  Liu C;
XX
DR   WPI; 2002-713376/77.
XX
PT   New isolated human nucleic acid molecule and polypeptide, useful for
PT   identifying, diagnosing, monitoring, staging, imaging and treating lung
PT   cancer and non-cancerous diseases of the lung.
XX
PS   Claim 1; Page 238-240; 389pp; English.
XX
CC   The invention describes an isolated human nucleic acid (I) encoding any
CC   of 120 10-1533 residue amino acid sequences (S1), given in the
CC   specification, comprising any of 164 179-12421 base pair sequences (S2),
CC   given in the specification. The methods and compositions of the present
CC   invention are useful for identifying, diagnosing, monitoring, staging,
CC   imaging and treating lung cancer and non-cancerous diseases of the lung.
CC   They are also used for identifying lung tissue, monitoring and
CC   identifying and/or designing antagonists of the polypeptide of the
CC   invention, gene therapy, production of transgenic animals and production
CC   of engineered lung tissue for treatment and research. This sequence
CC   encodes a lung specific nucleic acid
XX
SQ   Sequence 2399 BP; 631 A; 653 C; 622 G; 493 T; 0 U; 0 Other;

  Query Match             87.1%;  Score 2397.4;  DB 6;  Length 2399;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2398;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        356 TTGAGGGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGC 415
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db       2399 TTGAGGGCCAGTGAGGTGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGC 2340

Qy        416 ACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCT 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2339 ACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCT 2280

Qy        476 GGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGT 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2279 GGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGT 2220

Qy        536 GCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGG 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 GCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGG 2160

Qy        596 CAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGA 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2159 CAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGA 2100

Qy        656 TGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2099 TGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGG 2040

Qy        716 GGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTC 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2039 GGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTC 1980

Qy        776 GTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAA 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1979 GTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAA 1920

Qy        836 AGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGA 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1919 AGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGA 1860

Qy        896 CTGCCCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTG 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1859 CTGCCCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTG 1800

Qy        956 ATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAA 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1799 ATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAA 1740

Qy       1016 AAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTA 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1739 AAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTA 1680

Qy       1076 TTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1679 TTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATA 1620

Qy       1136 TTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAA 1195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1619 TTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAA 1560

Qy       1196 AAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAG 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1559 AAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAG 1500

Qy       1256 AAAGCCTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAA 1315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 AAAGCCTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAA 1440

Qy       1316 GCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTG 1375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1439 GCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTG 1380

Qy       1376 TCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGA 1435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 TCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGA 1320

Qy       1436 GTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTT 1495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 GTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTT 1260

Qy       1496 CCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGT 1555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 CCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGT 1200

Qy       1556 TGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCAC 1615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 TGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCAC 1140

Qy       1616 ACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGG 1675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 ACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGG 1080

Qy       1676 ACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGC 1735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 ACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGC 1020

Qy       1736 TACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGC 1795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 TACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGC 960

Qy       1796 CCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCC 1855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCC 900

Qy       1856 TCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCC 1915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 TCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCC 840
```

```
Qy       1916 CTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGC 1975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 CTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGC 780

Qy       1976 CACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAG 2035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 CACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAG 720

Qy       2036 GCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAG 2095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 GCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAG 660

Qy       2096 GCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCC 2155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 GCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCC 600

Qy       2156 ACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTG 2215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 ACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTG 540

Qy       2216 CAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCC 2275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCC 480

Qy       2276 TGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGG 2335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGG 420

Qy       2336 TTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGG 2395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGG 360

Qy       2396 TGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAG 2455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAG 300

Qy       2456 ATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTAT 2515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 ATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTAT 240

Qy       2516 CTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTG 2575
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTG 180

Qy       2576 GGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTT 2635
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 GGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTT 120

Qy       2636 GCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGT 2695
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 GCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGT 60

Qy       2696 TAGAATGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         59 TAGAATGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAAGATGCGGCC 1


RESULT 4
ABX92057
ID   ABX92057 standard; cDNA; 960 BP.
XX
AC   ABX92057;
XX
DT   08-MAY-2003  (first entry)
XX
DE   Lung specific nucleic acid (LSNA) #99.
XX
KW   Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;
KW   cancer monitoring; cancer staging; cancer imaging; lung cancer;
KW   non-cancerous diseases of the lung; transgenic animal; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200268633-A2.
XX
PD   06-SEP-2002.
XX
PF   21-NOV-2001; 2001WO-US043612.
XX
PR   22-NOV-2000; 2000US-0252500P.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Macina RA,  Recipon H,  Chen S,  Sun Y,  Liu C;
XX
DR   WPI; 2002-713376/77.
XX
PT   New isolated human nucleic acid molecule and polypeptide, useful for
PT   identifying, diagnosing, monitoring, staging, imaging and treating lung
PT   cancer and non-cancerous diseases of the lung.
XX
PS   Claim 1; Page 246; 389pp; English.
XX
CC   The invention describes an isolated human nucleic acid (I) encoding any
CC   of 120 10-1533 residue amino acid sequences (S1), given in the
CC   specification, comprising any of 164 179-12421 base pair sequences (S2),
CC   given in the specification. The methods and compositions of the present
CC   invention are useful for identifying, diagnosing, monitoring, staging,
CC   imaging and treating lung cancer and non-cancerous diseases of the lung.
CC   They are also used for identifying lung tissue, monitoring and
CC   identifying and/or designing antagonists of the polypeptide of the
CC   invention, gene therapy, production of transgenic animals and production
CC   of engineered lung tissue for treatment and research. This sequence
CC   encodes a lung specific nucleic acid
XX
SQ   Sequence 960 BP; 183 A; 351 C; 232 G; 193 T; 0 U; 1 Other;

  Query Match             22.2%;  Score 612.4;  DB 6;  Length 960;
  Best Local Similarity   99.4%;  Pred. No. 1e-156;
  Matches  646;  Conservative    0;  Mismatches    1;  Indels    3;  Gaps    3;

Qy       1658 GTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTG 1717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTG 100

Qy       1718 TCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTT 1777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 TCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTT 160

Qy       1778 GAGGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCC 1837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 GAGGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCC 220

Qy       1838 ATTGGAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGG 1897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 ATTGGAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGG 280

Qy       1898 CTGTGGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACT 1957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 CTGTGGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACT 340

Qy       1958 CCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAG 2017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 CCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAG 400

Qy       2018 CTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCT 2077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 CTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCT 460

Qy       2078 CCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTA 2137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTA 520

Qy       2138 GATGGGGGCCAGGGTCCCACA-GCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGT 2196
              ||||||||||||||||||||| ||| |||||||||||||||||||||||||||||||||
```

```
Db         521 GATGGGGGCCAGGGTCCCACATGCTGCTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGT 580

Qy        2197 CCAGTCCGGAT-CCCACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTG 2255
               ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         581 CCAGTCCGGATCCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTG 640

Qy        2256 TCACAGTGCTGG-CTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTC 2304
               |||||||||||| |||||||||||||||||||||||||||||||||||||
Db         641 TCACAGTGCTGGCCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTC 690


RESULT 5
ABV54530
ID   ABV54530 standard; cDNA; 611 BP.
XX
AC   ABV54530;
XX
DT   17-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 54521.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US005171.
XX
PR   17-FEB-2000; 2000US-0183319P.
PR   16-MAR-2000; 2000US-0189862P.
PR   25-MAY-2000; 2000US-0207454P.
PR   09-JUN-2000; 2000US-0211314P.
PR   18-JUL-2000; 2000US-0219007P.
PR   13-DEC-2000; 2000US-0255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS   Claim 1; Page 10539-10540; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for: (a) assessing whether
CC   a patient is afflicted with prostate cancer; (b) monitoring the
CC   progression of prostate cancer in a patient; (c) assessing the efficacy
CC   of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC   the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC   determining whether prostate cancer has metastasized in a patient; (h)
CC   assessing the aggressiveness or indolence of prostate cancer in a patient
CC   ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ   Sequence 611 BP; 111 A; 174 C; 164 G; 159 T; 0 U; 3 Other;

  Query Match             20.1%;  Score 554;  DB 5;  Length 611;
  Best Local Similarity   99.5%;  Pred. No. 7.8e-141;
  Matches  565;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy        1194 AAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAA 1253
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          44 AAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAA 103
```

```
Qy        1254 AGAAAGCCTCAGACAGTAAGGGC-TTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTG 1312
               ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db         104 AGAAAGCCTCAGACAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTG 163

Qy        1313 GAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTC 1372
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         164 GAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTC 223

Qy        1373 GTGTCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGC 1432
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         224 GTGTCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGC 283

Qy        1433 AGAGTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTG 1492
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         284 AGAGTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTG 343

Qy        1493 CTTCCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAG 1552
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         344 CTTCCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAG 403

Qy        1553 GGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCT 1612
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         404 GGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCT 463

Qy        1613 CACACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTG 1672
               ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         464 CACACTCTTGTCCTGNGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTG 523

Qy        1673 AGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCC 1732
               |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db         524 AGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTNCAGCTGCTCTTCC 583

Qy        1733 TGCTACAAAGGGGACTGCTCACAGTGGC 1760
               ||||||||||||||||||||||||||||
Db         584 TGCTACAAAGGGGACTGCTCACAGTGGC 611


RESULT 6
ABT07522
ID   ABT07522 standard; cDNA; 591 BP.
XX
AC   ABT07522;
XX
DT   14-NOV-2002  (first entry)
XX
DE   Human breast cancer associated coding sequence SEQ ID NO: 1.
XX
KW   Human; breast specific gene; breast specific protein; breast cancer;
KW   gene therapy; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200264611-A1.
XX
PD   22-AUG-2002.
XX
PF   12-FEB-2002; 2002WO-US004197.
XX
PR   13-FEB-2001; 2001US-0268292P.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Salceda S,  Macina RA,  Hu P,  Recipon H,  Karra K,  Cafferkey R;
PI   Sun Y,  Liu C;
XX
DR   WPI; 2002-657582/70.
XX
PT   New breast specific nucleic acids and proteins, useful for identifying,
PT   diagnosing, monitoring, staging, imaging, and treating breast cancer and
PT   non-cancerous disease states in breast tissue, and in gene therapy.
XX
PS   Claim 1; Page 148; 367pp; English.
```

```
XX
CC   The present invention provides human breast specific coding sequences and
CC   proteins. These can be used in the diagnosis and treatment of breast
CC   cancer and non-cancerous diseases of the breast. The present sequence is
CC   a coding sequence of the invention
XX
SQ   Sequence 591 BP; 106 A; 166 C; 172 G; 147 T; 0 U; 0 Other;

  Query Match             17.6%;  Score 483.8;  DB 6;  Length 591;
  Best Local Similarity   97.5%;  Pred. No. 1.3e-121;
  Matches  555;  Conservative    0;  Mismatches    7;  Indels    7;  Gaps    6;

Qy       1735 CTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGG 1794
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 CTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGG 71

Qy       1795 CCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGA-GGATGGACAG 1853
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||| ||||
Db         72 CCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGATGGATGTACAG 131

Qy       1854 CCTCAAATGGAAG-GAGTCCCACGGGAGATGGGTCCGAGGTCC-GGCTGTGGCCATCCAG 1911
              ||||||||||||| ||||||||||||||||||||||||||||| ||||||||||||||||
Db        132 CCTCAAATGGAAGTGAGTCCCACGGGAGATGGGTCCGAGGTCCAGGCTGTGGCCATCCAG 191

Qy       1912 CCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAG 1971
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAG 251

Qy       1972 CAGCCACTGCAGTTCCTTTCTTCGTGAG-TAACAGTAGTGATAGCAGCTGGGGCTAACAG 2030
               ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        252 TAGCCACTGCAGTTCCTTTCTTCGTGAGATAACAGTAGTGATAGCAGCTGGGGCTAACAG 311

Qy       2031 GCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATG 2090
              ||||||||| |||| |||||||||||||||||||||||||||||||||||||||||||||
Db        312 GCTAGGCTTAGTGTCCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATG 371

Qy       2091 GGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAG- 2149
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 GGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGA 431

Qy       2150 GGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTT--CCTGGTGTCCAGTCCGGAT 2207
              |||||||||||||||||||||||||||||||||||||||  |  | ||||||||||||||
Db        432 GGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTTCCTGGATGTCCAGTCCGGAT 491

Qy       2208 CCCACTTGCAGATCTCATGCTCTCAGATAGGTGGACAAGTTCTTTTGTCACAGTGCTGG 2267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 CCCACTTGCAGATCTCATGCTCTCAGATAGGTGGACAAGTTCTTTTGTCACAGTGCTGG 551

Qy       2268 CTCTGTCCTGAGGCCTCATTGCTGGCTGG 2296
              |||||||||||||||||||||||||||||
Db        552 CTCTGTCCTGAGGCCTCATTGCTGGCTGG 580


RESULT 7
AAS59058
ID   AAS59058 standard; cDNA; 395 BP.
XX
AC   AAS59058;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Human cancer related cDNA sequence #176.
XX
KW   Human; ss; lung cancer; adenocarcinoma; breast cancer; colon cancer;
KW   prostate cancer; benign prostatic hypertrophy; BHP; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200172781-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US009952.
XX
PR   28-MAR-2000; 2000US-0192583P.
XX
PA   (CHIR ) CHIRON CORP.
PA   (HYSE-) HYSEQ INC.
XX
PI   Williams LT,  Escobedo J,  Innis MA,  Garcia PD,  Sudduth-Klinger J;
PI   Reinhard C,  He Z,  Randazzo F,  Kennedy GC,  Pot D,  Kassam A;
PI   Lamson G,  Drmanac R,  Crkvenjakov R,  Dickson M,  Drmanac S,  Labat I;
PI   Leshkowitz D,  Kita D,  Garcia V,  Jones LW,  Stache-Crain B;
XX
DR   WPI; 2001-626251/72.
XX
PT   New human polynucleotides useful for the treatment and diagnosis of
PT   cancer.
XX
PS   Claim 1; Page 202; 240pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising a
CC   nucleotide sequence which hybridises to a sequence selected from one of
CC   316 fully defined sequences given in the specification, antisense
CC   molecules complimentary to the sequences, the polypeptides encoded by the
CC   sequences and antibodies raised against the proteins. The nucleic acids
CC   are useful for detecting differentially expressed genes which correlate
CC   with a cancerous state of a mammalian cell i.e. diagnosing cancer
CC   (especially lung cancer, colon cancer, breast cancer, prostate cancer and
CC   adenocarcinoma). Modifying the gene products of of the nucleic acids
CC   results in inhibition of tumour growth. The nucleic acids are also useful
CC   in gene mapping and tissue profiling. The present sequence is one of the
CC   316 cancer related cDNA sequences
XX
SQ   Sequence 395 BP; 70 A; 111 C; 109 G; 105 T; 0 U; 0 Other;

  Query Match             13.0%;  Score 357.4;  DB 4;  Length 395;
  Best Local Similarity   96.9%;  Pred. No. 4.4e-87;
  Matches  375;  Conservative    0;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       1216 CAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGG 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 CAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGG 68

Qy       1276 C-TTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGAT 1334
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CTTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGAT 128

Qy       1335 GAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGC 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 GAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGC 188

Qy       1395 CCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGT 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 CCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGT 248

Qy       1455 GGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGC 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 GGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGC 308

Qy       1515 TGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACT 1574
              |||||||||||||||||||||||||||||||||||||||| ||||||| | |||||||||
Db        309 TGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGCTGATGCGCGTAGAAGGACT 368

Qy       1575 TTGAGCAGTGGTGGGCAGCAGTGGCCT 1601
              ||||||||||||||       || |||
Db        369 TTGAGCAGTGGTGGCAGCACTTGCCCT 395


RESULT 8
ADD49697
ID   ADD49697 standard; cDNA; 295 BP.
XX
AC   ADD49697;
XX
DT   15-JAN-2004  (first entry)
```

```
XX
DE   Human lung cancer associated cDNA 61594667.
XX
KW   Human; ss; lung cancer antigen; cytostatic; lung cancer; gene therapy;
KW   vaccine; T-cell; tumour.
XX
OS   Homo sapiens.
XX
PN   US2003194764-A1.
XX
PD   16-OCT-2003.
XX
PF   04-APR-2002; 2002US-00116712.
XX
PR   05-APR-2001; 2001US-0282289P.
PR   05-OCT-2001; 2001US-0327511P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Bangur CS,  Switzer A;
XX
DR   WPI; 2003-844452/78.
XX
PT   New isolated polypeptides and polynucleotides useful for diagnosing,
PT   preventing and treating cancer, particularly lung cancer.
XX
PS   Claim 1; SEQ ID NO 429; 250pp; English.
XX
CC   The invention relates to an isolated polynucleotide (a) comprising any of
CC   the 666 fully defined nucleotide sequences appearing as ADD49269 -
CC   ADD49935, ADD49936 and ADD49938, complements of (a); sequences of at
CC   least 20 contiguous residues of (a); sequences that hybridise to (a)
CC   under highly stringent conditions; sequences having at least 75 or 90%
CC   identity to (a); or degenerate variants of (a). Also included are an
CC   isolated polypeptide ((b)(comprising: sequences encoded by the new
CC   polynucleotide; any of the 4 amino acid sequences fully defined in the
CC   specification; or sequences having at least 70 or 90% identity to the
CC   sequence in (a) or (b)), an expression vector comprising the above
CC   polynucleotide operably linked to an expression control sequence, a host
CC   cell transformed or transfected with the above expression vector, an
CC   isolated antibody, or its antigen-binding fragment, that specifically
CC   binds to the above polypeptide, an oligonucleotide that hybridises to the
CC   above-mentioned nucleotide sequences under highly stringent conditions, a
CC   fusion protein comprising at least one polypeptide cited above, detecting
CC   the presence of a cancer in a patient (comprising: obtaining a biological
CC   sample from the patient; contacting the biological sample with a binding
CC   agent that binds to the polypeptide, or with the oligonucleotide cited
CC   above; detecting in the sample an amount of the polypeptide that binds to
CC   the binding agent, or an amount of a polynucleotide that hybridises to
CC   the oligonucleotide; and comparing the amount of polypeptide, or
CC   polynucleotide that hybridises to the oligonucleotide, to a predetermined
CC   cut-off value and then determining the presence of a cancer in the
CC   patient), a method for stimulating and/or expanding T-cells specific for
CC   a tumour protein (comprising contacting T-cells with the above
CC   polypeptide, polynucleotide or antigen-presenting cells that express the
CC   polynucleotide, under conditions and for a time sufficient to permit the
CC   stimulation and/or expansion of T-cells), an isolated T-cell population
CC   comprising T-cells prepared by the method, a composition comprising a
CC   first component selected from physiological carriers and
CC   immunostimulants, and a second component selected from the above
CC   polypeptide, polynucleotide, antibody, fusion protein, T-cell population
CC   and antigen-presenting cells that express the above polypeptide,
CC   stimulating an immune response in a patient (comprising administering to
CC   the patient the above composition) treating lung cancer in a patient
CC   (comprising administering to the patient the above composition and a
CC   diagnostic kit (comprising: at least one oligonucleotide cited above; or
CC   at least one antibody cited above and a detection reagent, where the
CC   detection reagent comprises a reporter group). The composition and
CC   methods are useful in diagnosing, preventing and treating cancer,
CC   particularly lung cancer. The present sequence is a lung cancer-
CC   associated antigen cDNA of the invention.
XX
SQ   Sequence 295 BP; 52 A; 88 C; 88 G; 67 T; 0 U; 0 Other;
```

```
  Query Match             10.7%;  Score 295;  DB 10;  Length 295;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-70;
  Matches  295;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1723 CTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGG 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGG 60

Qy       1783 GCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGG 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGG 120

Qy       1843 AGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTG 1902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTG 180

Qy       1903 GCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGG 1962
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGG 240

Qy       1963 GGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAG 2017
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAG 295


RESULT 9
ABX92048/c
ID   ABX92048 standard; cDNA; 251 BP.
XX
AC   ABX92048;
XX
DT   08-MAY-2003  (first entry)
XX
DE   Lung specific nucleic acid (LSNA) #90.
XX
KW   Human; gene therapy; vaccine; lung specific antigen; cancer diagnosis;
KW   cancer monitoring; cancer staging; cancer imaging; lung cancer;
KW   non-cancerous diseases of the lung; transgenic animal; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200268633-A2.
XX
PD   06-SEP-2002.
XX
PF   21-NOV-2001; 2001WO-US043612.
XX
PR   22-NOV-2000; 2000US-0252500P.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Macina RA,  Recipon H,  Chen S,  Sun Y,  Liu C;
XX
DR   WPI; 2002-713376/77.
XX
PT   New isolated human nucleic acid molecule and polypeptide, useful for
PT   identifying, diagnosing, monitoring, staging, imaging and treating lung
PT   cancer and non-cancerous diseases of the lung.
XX
PS   Claim 1; Page 238; 389pp; English.
XX
CC   The invention describes an isolated human nucleic acid (I) encoding any
CC   of 120 10-1533 residue amino acid sequences (S1), given in the
CC   specification, comprising any of 164 179-12421 base pair sequences (S2),
CC   given in the specification. The methods and compositions of the present
CC   invention are useful for identifying, diagnosing, monitoring, staging,
CC   imaging and treating lung cancer and non-cancerous diseases of the lung.
CC   They are also used for identifying lung tissue, monitoring and
CC   identifying and/or designing antagonists of the polypeptide of the
CC   invention, gene therapy, production of transgenic animals and production
CC   of engineered lung tissue for treatment and research. This sequence
CC   encodes a lung specific nucleic acid
```

```
XX
SQ   Sequence 251 BP; 79 A; 75 C; 54 G; 43 T; 0 U; 0 Other;

  Query Match             8.1%;  Score 224.4;  DB 6;  Length 251;
  Best Local Similarity   98.8%;  Pred. No. 9.4e-51;
  Matches  237;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy       2381 CCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAAT 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 CCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAAT 192

Qy       2441 GATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAA 2500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAA 132

Qy       2501 AGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGA 2560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 AGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGA 72

Qy       2561 GGTCCTGGGGAGGT--GGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCT 2618
              ||||||||||||||  | ||||||||||||||||||||||||||||||||||||||||||
Db         71 GGTCCTGGGGAGGTGCGCGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCT 12


RESULT 10
AAZ35057
ID   AAZ35057 standard; DNA; 736 BP.
XX
AC   AAZ35057;
XX
DT   28-FEB-2000  (first entry)
XX
DE   Pluripotent cell marker gene L17 DD-PCR product.
XX
KW   Cell differentiation; cell proliferation; cell maintenance; human;
KW   ectoderm-like cell; embryonic stem cell; pluripotent cell; gene therapy;
KW   cell therapy; tissue transplant; organ transplant; xerotransplant;
KW   allotransplant; concomitant transplantation; transgenic animal; L17; ds.
XX
OS   Homo sapiens.
XX
PN   WO9953021-A1.
XX
PD   21-OCT-1999.
XX
PF   09-APR-1999;   99WO-AU000265.
XX
PR   09-APR-1998;   98AU-00002912.
PR   23-SEP-1998;   98AU-00006097.
XX
PA   (BRES-) BRESAGEN LTD.
XX
PI   Bettess MD,  Rathjen PD,  Rathjen J;
XX
DR   WPI; 2000-061970/05.
XX
PT   New isolated biologically active factor capable of influencing
PT   differentiation, proliferation or maintenance of pluripotent cells.
XX
PS   Example 1; Fig 3A; 189pp; English.
XX
CC   This is the nucleotide sequence of a novel partial cDNA, termed L17, that
CC   is differentially expressed between embryonic stem (ES) and primitive
CC   ectoderm-like (EPL) cells. Novel L17, K7 (see AAZ35058) and Psc1 (see
CC   AAZ35059) clones were identified by differential display PCR using RNA
CC   isolated form ES and EPL cells that had been cultured in MEDII+LIF for 2,
CC   4, 6 and 8 days. L17 was highly expressed in ES cells but down regulated
CC   rapidly with EPL formation and maintenance in culture. The expression of
CC   novel pluripotent cell marker genes supports the relationship between EPL
CC   cells and early primitive ectoderm. The invention relates to a novel
CC   biologically active factor capable of influencing the differentiation,
CC   proliferation or maintenance of pluripotent cells. Methods are provided
CC   for producing partially or terminally differentiated cells from
```

```
CC   pluripotent cells. The pluripotent cells and partially or terminally
CC   differentiated cells can be used in allo-, concomitant- or xeno-
CC   transplantation, cell therapy, tissue and organ augmentation or
CC   replacement, and gene therapy. They can also be used for producing
CC   chimeric or transgenic animals
XX
SQ   Sequence 736 BP; 202 A; 147 C; 174 G; 213 T; 0 U; 0 Other;

  Query Match             2.8%;  Score 76.6;  DB 3;  Length 736;
  Best Local Similarity   68.5%;  Pred. No. 5.4e-10;
  Matches  191;  Conservative    0;  Mismatches   59;  Indels   29;  Gaps    5;

Qy       2471 CTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGT 2530
              ||| ||| || | |||||| ||||||||||||||||| ||  | | |    ||  |
Db        483 CTTCCTCTAGGTGGGGTGATTCTTTTCTAAAGGTGGTAGGGTTTTTTAGGTGGTAG---- 538

Qy       2531 GGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTC-CTGGGGAGGTGGGCGTTGAGCTCAA 2589
                      ||  |||||  |  |||||  ||||| | ||  |||  | | | ||   | |
Db        539 ----TCCAGACTGGTCAAATATGTGTGGCAGGTCACGGGAAAGGGAGTCATGGAAACCCA 594

Qy       2590 AGTTGTCCTACTGCCA---TGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTTAC 2646
               |||||| || |||||   |||||||||||||||||||||||||||||||||||||| |
Db        595 CGTTGTCATAATGCCATGTTGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTAAA 654

Qy       2647 TGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAATGCAAA 2706
              || |                || |||||||||||| ||||||| ||||| ||||||||||
Db        655 TGAA----------------AAATCTGTATGTGGGCTCTGTGC-TGGGTCAGAATGCAAA 697

Qy       2707 TAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAA 2745
              ||||||      |  ||| |||||||||||||| |||||
Db        698 TAAAACCATTTGTACTAAAAAAAAAAAAAAAAAAAAAAA 736


RESULT 11
AAK90877/c
ID   AAK90877 standard; DNA; 32038 BP.
XX
AC   AAK90877;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human digestive system antigen genomic sequence SEQ ID NO: 4453.
XX
KW   Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW   ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW   digestive system disorder; Meckel's diverticulum; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155314-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001324.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
```

```
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
```

```
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-502630/55.
XX
PT   Polynucleotides encoding digestive system antigens, useful for
PT   diagnosing, treating, preventing and/or prognozing disorders of the
PT   digestive system, particularly cancer and cancer metastases.
XX
PS   Disclosure; SEQ ID NO 4453; 986pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human digestive system antigens. These can be used in the
CC   diagnosis, treatment and prevention of digestive system disorders,
CC   including cancer, Meckel's diverticulum, bacterial or parasitic
CC   infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC   ulcerative colitis. The present sequence is a genomic DNA fragment
CC   encoding a digestive system antigen of the invention
XX
SQ   Sequence 32038 BP; 9776 A; 5476 C; 5963 G; 10823 T; 0 U; 0 Other;

  Query Match              2.7%;  Score 75;  DB 4;  Length 32038;
  Best Local Similarity   65.9%;  Pred. No. 1.3e-08;
  Matches  147;  Conservative    0;  Mismatches   60;  Indels   16;  Gaps    2;

Qy       1013 AAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATT 1072
```

```
               ||| |   ||||||| |  ||||||  ||  |||||    |||||  ||| ||  |||||||
Db       28877 AAATATTTAATTCTTTTCATTAGTAAACCCGTATTACAAAAAAAATTGTTCTAAATTATT 28818

Qy        1073 GTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGC------- 1125
               |  ||||| ||||| |||||  ||   | ||| ||||||  || ||||   |
Db       28817 G--TTTTGAATTTTGTCAATAAAATTTTCGTGCAAATTTTATTTCTCTCTAGTTATTTAA 28760

Qy        1126 -------CAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGT 1178
                      ||  | ||| |   |||| |||||| || |  |||||| |||||||||||  |
Db       28759 GTATCTACATAACAATCTCAGTTTTACCTCTTAGCCCACAAAGCCTAAAATATTTACTAT 28700

Qy        1179 CTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
               || |||| ||| ||||||||| | |||||  |||  | ||| |
Db       28699 CTGGCCCTTTATAGAAAAAGTTTCCTGACCTCTGCCCTAGAGC 28657


RESULT 12
AAK89980/c
ID   AAK89980 standard; DNA; 32038 BP.
XX
AC   AAK89980;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human digestive system antigen genomic sequence SEQ ID NO: 3556.
XX
KW   Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW   ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW   digestive system disorder; Meckel's diverticulum; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155314-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001324.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
```

```
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-502630/55.
XX
PT   Polynucleotides encoding digestive system antigens, useful for
PT   diagnosing, treating, preventing and/or prognozing disorders of the
PT   digestive system, particularly cancer and cancer metastases.
XX
PS   Disclosure; SEQ ID NO 3556; 986pp; English.
XX
CC   The present invention provides the protein and coding sequences of a
CC   number of human digestive system antigens. These can be used in the
CC   diagnosis, treatment and prevention of digestive system disorders,
CC   including cancer, Meckel's diverticulum, bacterial or parasitic
CC   infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC   ulcerative colitis. The present sequence is a genomic DNA fragment
CC   encoding a digestive system antigen of the invention
XX
SQ   Sequence 32038 BP; 9776 A; 5476 C; 5963 G; 10823 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 75;  DB 4;  Length 32038;
  Best Local Similarity   65.9%;  Pred. No. 1.3e-08;
  Matches  147;  Conservative    0;  Mismatches   60;  Indels   16;  Gaps    2;

Qy       1013 AAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATT 1072
              ||| |   ||||||| |  |||||| ||  |||||    |||||  ||| || |||||||
Db      28877 AAATATTTAATTCTTTTCATTAGTAAACCCGTATTACAAAAAAAATTGTTCTAAATTATT 28818

Qy       1073 GTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGC------- 1125
              |  ||||| ||||| |||||  ||   | ||| ||||||  || ||||   |
Db      28817 G--TTTTGAATTTTGTCAATAAAATTTTCGTGCAAATTTTATTTCTCTCTAGTTATTTAA 28760

Qy       1126 -------CAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGT 1178
                     ||  | ||| |   |||| |||||| || |  |||||| |||||||||||  |
Db      28759 GTATCTACATAACAATCTCAGTTTTACCTCTTAGCCCACAAAGCCTAAAATATTTACTAT 28700

Qy       1179 CTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
              || |||| ||| ||||||||| | |||||  |||  | ||| |
Db      28699 CTGGCCCTTTATAGAAAAAGTTTCCTGACCTCTGCCCTAGAGC 28657


RESULT 13
```

```
AAS30028/c
ID   AAS30028 standard; DNA; 32038 BP.
XX
AC   AAS30028;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human lung antigen genomic DNA #98.
XX
KW   Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;
KW   chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW   cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW   ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW   hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW   cerebrovascular disorder; nervous system disorder; bacterial infection;
KW   fungal infection; viral infection; ocular disorder; endocrine disorder;
KW   gastrointestinal disorder; renal disorder; respiratory disorder;
KW   wound healing; skin aging; organ transplantation; food preservative;
KW   tissue regeneration; anti-infertility; food additive.
XX
OS   Homo sapiens.
XX
PN   WO200155303-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001301.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
```

```
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-457723/49.
XX
PT   Isolated polypeptide for treating, preventing and/ or prognosing
PT   respiratory disorders related to the lung including lung cancers and also
PT   for testing and detection e.g. diagnosis.
XX
PS   Claim 1; SEQ ID NO 292; 507pp; English.
XX
CC   Sequences AAS29931-AAS30164 represent genomic DNA molecules, which encode
CC   the lung antigen polypeptides of the invention. Lung antigen polypeptides
CC   and their associated polynucleotides are useful in the diagnosis,
CC   treatment and prevention of various types of disorders in e.g. humans,
CC   mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
CC   pathological condition can be determined by detecting the presence or
CC   absence of a mutation in a lung antigen polynucleotide. The treatable
CC   disorders include autoimmune diseases such as rheumatoid arthritis,
CC   hyperproliferative disorders such as neoplasms of the breast or liver,
CC   cardiovascular disorders such as cardiac arrest, cerebrovascular
CC   disorders such as cerebral ischaemia, nervous system disorders such as
CC   Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC   ocular disorders such as corneal infection, endocrine disorders such as
CC   premature labour and infertility, gastrointestinal disorders such as
CC   Crohn's disease, renal disorders such as glomerulonephritis and
CC   respiratory disorders such as asthma and pleurisy. The polypeptides can
CC   also be used to aid wound healing, to prevent skin aging due to sunburn,
CC   to maintain organs before transplantation, to regenerate tissues and in
CC   chemotaxis. The polypeptides can also be used as a food additive or
CC   preservative to increase or decrease storage capabilities. Note: The
CC   sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences

  Query Match             2.7%;  Score 75;  DB 5;  Length 32038;
  Best Local Similarity   65.9%;  Pred. No. 1.3e-08;
  Matches  147;  Conservative    0;  Mismatches   60;  Indels   16;  Gaps    2;

Qy       1013 AAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATT 1072
              ||| |   ||||||| |  |||||| ||  |||||    |||||  ||| || |||||||
Db      28877 AAATATTTAATTCTTTTCATTAGTAAACCCGTATTACAAAAAAAATTGTTCTAAATTATT 28818

Qy       1073 GTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGC------- 1125
              |  ||||| ||||| |||||  ||   | ||| ||||||  || ||||   |
Db      28817 G--TTTTGAATTTTGTCAATAAAATTTTCGTGCAAATTTTATTTCTCTCTAGTTATTTAA 28760

Qy       1126 -------CAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGT 1178
                     ||  | ||| |   |||| |||||| || |  |||||| |||||||||||  |
Db      28759 GTATCTACATAACAATCTCAGTTTTACCTCTTAGCCCACAAAGCCTAAAATATTTACTAT 28700

Qy       1179 CTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
```

```
              || |||| ||| ||||||||| | |||||  |||  | ||| |
Db      28699 CTGGCCCTTTATAGAAAAAGTTTCCTGACCTCTGCCCTAGAGC 28657


RESULT 14
ADB33365/c
ID   ADB33365 standard; DNA; 32038 BP.
XX
AC   ADB33365;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human novel lung related polypeptide DNA SEQ ID NO 292.
XX
KW   gene therapy; lung antigen; neoplasia; acute myelogenous leukaemia;
KW   adenocarcinoma; respiratory disorder; chronic rhinitis; sinusitis;
KW   immunodeficiency; X-linked agammaglobulinaemia;
KW   X-linked infantile agammaglobulinaemia; inflammatory disorder;
KW   adrenalitis; alveolitis; immune complex disease; serum sickness;
KW   polyarteritis nodosa; bleeding disorder; thrombocytopenia;
KW   Von Willebrand's disease; acquired platelet dysfunction; kidney failure;
KW   multiple myeloma; macrophage related disorder; Gaucher's disease;
KW   Neimann-Pick disease; tumour; colon cancer; pancreatic cancer;
KW   renal disorder; nephritis; bone disorder; Albers-Schonberg disease;
KW   bowleg; muscle disorder; Becker's muscular dystrophy;
KW   Duchenne's muscular dystrophy; nervous disorder; ischaemic lesion;
KW   traumatic lesion; endocrine disorder; Cushing's syndrome;
KW   corticosteroid deficiency; gastrointestinal disorder; dysphagia;
KW   gastric reflux; human; ds.
XX
OS   Homo sapiens.
XX
PN   US2003054368-A1.
XX
PD   20-MAR-2003.
XX
PF   22-FEB-2002; 2002US-00079854.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
```

```
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
PR   17-JAN-2001; 2001US-00764878.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Barash SC;
XX
DR   WPI; 2003-695900/66.
XX
PT   Novel isolated lung antigen polypeptides useful for treating, preventing,
PT   diagnosing acute myelogenous leukemias, adenocarcinoma, thrombocytopenia,
PT   Von Willebrand's disease.
XX
PS   Disclosure; SEQ ID NO 292; 178pp; English.
XX
CC   The invention relates to an isolated lung antigen polypeptide sequence or
CC   encoded sequence in a cDNA clone. The polypeptide and its polynucleotide
CC   are useful for treating, preventing, diagnosing and/or prognosing
CC   diseases and/or disorders such as pathological cell proliferative
CC   neoplasias e.g. acute myelogenous leukaemias, adenocarcinoma; respiratory
CC   disorders such as chronic rhinitis, sinusitis; immunodeficiencies such as
CC   X-linked agammaglobulinaemia, X-linked infantile agammaglobulinaemia;
CC   inflammatory disorders such as adrenalitis, alveolitis; immune complex
CC   diseases such as serum sickness, polyarteritis nodosa; bleeding disorders
CC   such as thrombocytopenia, Von Willebrand's disease; acquired platelet
CC   dysfunction such as kidney failure, multiple myeloma; disorders
CC   associated with macrophage numbers and/or macrophage function such as
CC   Gaucher's disease, Neimann-Pick disease; tumours such as colon cancer,
CC   pancreatic cancer; renal disorders such as kidney failure, nephritis;
CC   bone disorders such as Albers-Schonberg disease, bowlegs; muscle
CC   disorders such as Becker's muscular dystrophy, Duchenne's muscular
CC   dystrophy; nervous disorders such as ischaemic lesions, traumatic lesions
CC   ; endocrine disorders such as Cushing's syndrome, corticosteroid

  Query Match             2.7%;  Score 75;  DB 10;  Length 32038;
  Best Local Similarity   65.9%;  Pred. No. 1.3e-08;
  Matches  147;  Conservative    0;  Mismatches   60;  Indels   16;  Gaps    2;

Qy       1013 AAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATT 1072
              ||| |   ||||||| |  |||||| ||  |||||    |||||  ||| ||  |||||||
Db      28877 AAATATTTAATTCTTTTCATTAGTAAACCCGTATTACAAAAAAAATTGTTCTAAATTATT 28818

Qy       1073 GTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGC------- 1125
              |  ||||| ||||| |||||  ||   | ||| ||||||  || ||||   |
Db      28817 G--TTTTGAATTTTGTCAATAAAATTTTCGTGCAAATTTTATTTCTCTCTAGTTATTTAA 28760

Qy       1126 -------CAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGT 1178
                     ||  | ||| |   |||| |||||| || |  |||||| |||||||||||  |
Db      28759 GTATCTACATAACAATCTCAGTTTTACCTCTTAGCCCACAAAGCCTAAAATATTTACTAT 28700

Qy       1179 CTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
              || |||| ||| ||||||||| | |||||  |||  | ||| |
Db      28699 CTGGCCCTTTATAGAAAAAGTTTCCTGACCTCTGCCCTAGAGC 28657


RESULT 15
ADJ12644
ID   ADJ12644 standard; DNA; 32844 BP.
XX
AC   ADJ12644;
XX
DT   20-MAY-2004  (first entry)
XX
DE   DNA fragment of a BAC clone that encodes a human secreted protein Seq498.
XX
KW   human; secreted; cancer; haematopoietic disease; anaemia;
KW   multiple myeloma; reproductive system disorder; prostatitis;
KW   inguinal hernia; musculoskeletal disease; systemic lupus erythematosus;
KW   gout; cardiovascular disease; arrhythmia; hypernatraemia; fetal disease;
KW   fetal alcohol syndrome; Down's syndrome; excretory disease;
KW   urinary incontinence; renal disorder; neural; sensory disease;
KW   Alzheimer's disease; meningitis; respiratory disease; emphysema;
KW   occupational lung disease; endocrine disease; diabetes;
KW   glomerulonephritis; digestive disease; portal hypertension;
KW   irritable bowel syndrome; epithelial disease; scleroderma;
KW   epidermolysis bullosa; cytostatic; antianemic; antiarthritic;
KW   antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;
KW   antipsoriatic; antibacterial; osteopathic; dermatological; antigout;
KW   immunomodulator; antiarrhythmic; cardiant; nootropic; antilipemic;
KW   nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;
KW   antidiabetic; anabolic; hypertensive; vulnerary; ds.
XX
OS   Homo sapiens.
XX
PN   US2004010132-A1.
XX
PD   15-JAN-2004.
XX
PF   30-OCT-2001; 2001US-00984429.
XX
PR   09-OCT-1997;   97US-0061463P.
PR   09-OCT-1997;   97US-0061527P.
PR   09-OCT-1997;   97US-0061529P.
PR   09-OCT-1997;   97US-0061532P.
PR   09-OCT-1997;   97US-0061536P.
PR   09-OCT-1997;   97US-0071498P.
PR   08-OCT-1998;   98WO-US021142.
PR   08-APR-1999;   99US-00288143.
PR   01-NOV-2000; 2000US-0244591P.
XX
PA   (ROSE/) ROSEN C A.
PA   (BREW/) BREWER L A.
PA   (DUAN/) DUAN R D.
PA   (RUBE/) RUBEN S M.
PA   (FLOR/) FLORENCE K A.
PA   (GREE/) GREENE J M.
PA   (YOUN/) YOUNG P E.
PA   (FERR/) FERRIE A M.
PA   (YUGG/) YU G.
PA   (FLOR/) FLORENCE C.
PA   (EBNE/) EBNER R.
PA   (OLSE/) OLSEN H.
XX
PI   Rosen CA,  Brewer LA,  Duan RD,  Ruben SM,  Florence KA,  Greene JM;
PI   Young PE,  Ferrie AM,  Yu G,  Florence C,  Ebner R,  Olsen H;
XX
DR   WPI; 2004-090518/09.
XX
PT   New isolated nucleic acids and polypeptides, useful for diagnosing,
```

```
PT   treating, preventing or ameliorating diseases or disorders e.g. cancer,
PT   anemia, arthritis, asthma, inflammatory bowel disease or Alzheimer's
PT   disease.
XX
PS   Disclosure; SEQ ID NO 498; 286pp; English.
XX
CC   This invention relates to novel polynucleotides encoding human secreted
CC   proteins. Specifically, it refers to the vectors, host cells, recombinant
CC   and synthetic methods for producing human polynucleotides, polypeptides
CC   and antibodies. Furthermore, it relates to screening methods to identify
CC   agonists and antagonists that can be used to inhibit or enhance the
CC   production and function of the secreted proteins. The present invention
CC   describes these compositions as useful for diagnosing, treating or
CC   preventing disorders such as cancer, haematopoietic diseases including
CC   anaemia and multiple myeloma, reproductive system disorders including
CC   prostatitis and inguinal hernia, musculoskeletal diseases including
CC   systemic lupus erythematosus and gout, cardiovascular disease including
CC   arrhythmia and hypernatraemia, mixed fetal diseases including fetal
CC   alcohol syndrome and Down's syndrome, excretory diseases including
CC   urinary incontinence and renal disorders, neural or sensory disease
CC   including Alzheimer's disease and meningitis, respiratory disease
CC   including emphysema and occupational lung disease, endocrine diseases
CC   including diabetes and glomerulonephritis, digestive diseases including
CC   portal hypertension and irritable bowel syndrome and connective tissue or
CC   epithelial diseases including scleroderma and epidermolysis bullosa. As
CC   such, there are various activites such as cytostatic, antianemic,
CC   antiarthritic, antiasthmatic, anti-HIV, immunosuppressive,
CC   antiinflammatory, antipsoriatic, antibacterial, osteopathic,
CC   dermatological, antigout, immunomodulator, antiarrhythmic, cardiant,
CC   nootropic, antilipemic, nephrotropic, uropathic, neuroprotective,
CC   antiparkinsonian, tranquilizer, antidiabetic, anabolic, hypertensive and
CC   vulnerary. This polynucleotide is a DNA fragment of a BAC clone that
CC   encodes a human secreted protein of the invention. NOTE: This sequence
CC   does not appear in the printed specification but has been obtained in
CC   electronic format from the US patent office at the following web site
CC   www.seqdata.uspto.gov/sequence.html; Document ID: 20040010132.
XX
SQ   Sequence 32844 BP; 11125 A; 6089 C; 5587 G; 10043 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 75;  DB 12;  Length 32844;
  Best Local Similarity   65.9%;  Pred. No. 1.3e-08;
  Matches  147;  Conservative    0;  Mismatches   60;  Indels   16;  Gaps    2;

Qy       1013 AAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATT 1072
              ||| |   ||||||| |  |||||| ||  |||||    |||||  ||| || |||||||
Db       3968 AAATATTTAATTCTTTTCATTAGTAAACCCGTATTACAAAAAAAATTGTTCTAAATTATT 4027

Qy       1073 GTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGC------- 1125
              |  ||||| ||||| |||||  ||   | ||| ||||||  || ||||   |
Db       4028 G--TTTTGAATTTTGTCAATAAAATTTTCGTGCAAATTTTATTTCTCTCTAGTTATTTAA 4085

Qy       1126 -------CAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGT 1178
                     ||  | ||| |   |||| |||||| || |  |||||| |||||||||||  |
Db       4086 GTATCTACATAACAATCTCAGTTTTACCTCTTAGCCCACAAAGCCTAAAATATTTACTAT 4145

Qy       1179 CTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACC 1221
              || |||| ||| ||||||||| | |||||  |||  | ||| |
Db       4146 CTGGCCCTTTATAGAAAAAGTTTCCTGACCTCTGCCCTAGAGC 4188

Search completed: November 29, 2004, 14:40:49
Job time : 1294 secs
```

```
                     GenCore version 5.1.6
              Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:          November 29, 2004, 08:42:07 ; Search time 11585 Seconds
                                               (without alignments)
                                               11241.764 Million cell updates/sec

Title:           US-09-989-920-100
Perfect score:   2754
Sequence:        1 gccagaagcagcctcagctt..........aaaaataaaaagatgcggcc 2754

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters:      9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                     %
Result               Query
  No.     Score      Match Length DB  ID                        Description
-------------------------------------------------------------------------------
    1      2754      100.0    2754  6  AX535073                  AX535073 Sequence
c   2    2718.4       98.7  211305  9  AC079988                  AC079988 Homo sapi
c   3    2718.4       98.7  225203  2  AC140107                  AC140107 Homo sapi
c   4    2397.4       87.1    2399  6  AX535064                  AX535064 Sequence
    5    1747.6       63.5    1853  6  AR204690                  AR204690 Sequence
c   6     820.6       29.8   88036  2  AC139359                  AC139359 Homo sapi
    7       803       29.2   88036  2  AC139359                  AC139359 Homo sapi
    8     612.4       22.2     960  6  AX535072                  AX535072 Sequence
    9       554       20.1     611  6  CQ522682                  CQ522682 Sequence
   10     357.4       13.0     395  6  AX262096                  AX262096 Sequence
c  11     261.2        9.5     406 11   G38425                    G38425 SHGC-58087
c  12     224.4        8.1     251  6  AX535063                  AX535063 Sequence
c  13      87.6        3.2  190379  2  AC134136                  AC134136 Rattus no
   14      87.6        3.2  231580  2  AC136817                  AC136817 Rattus no
   15        81        2.9  182643  9  CNS01DT7                  AL132642 Human chr
c  16      80.6        2.9  184043  2  AC090296                  AC090296 Homo sapi
   17      80.6        2.9  188789  2  AP001446                  AP001446 Homo sapi
   18      80.6        2.9  211466  9  AC012301                  AC012301 Homo sapi
c  19        77        2.8  175762  2  AC009761                  AC009761 Homo sapi
   20        77        2.8  180760  2  AC130435                  AC130435 Homo sapi
c  21        77        2.8  194955  9  AC084198                  AC084198 Homo sapi
   22      76.6        2.8     736  6  BD274467                  BD274467 Cell diff
c  23      76.2        2.8  155301  9  AC041040                  AC041040 Homo sapi
   24      75.4        2.7  151687  2  AL356979                  AL356979 Homo sapi
   25      75.4        2.7  182529  9  AC092813                  AC092813 Homo sapi
c  26      75.4        2.7  202402  2  AC012517                  AC012517 Homo sapi
   27        75        2.7   65543  2  AC027356                  AC027356 Homo sapi
c  28        75        2.7  126882  9  HS273N12                  AL022395 Human DNA
c  29      74.2        2.7  181504  2  AC150610                  AC150610 Callithri
   30      73.2        2.7  125066  9  AC074286                  AC074286 Homo sapi
c  31      72.4        2.6  137658  9  HS1022J11                 AL049765 Human DNA
c  32      72.4        2.6  202157  4  AC124907                  AC124907 Equus cab
   33      71.6        2.6   56748  9  AL356852                  AL356852 Human DNA
c  34      71.6        2.6  169733  2  AC068046                  AC068046 Homo sapi
c  35      71.4        2.6  134942  9  AC008660                  AC008660 Homo sapi
c  36      70.8        2.6  123395  9  HS889N15                  AL031177 Human DNA
c  37      70.2        2.5  195634  2  AC150472                  AC150472 Papio anu
c  38        70        2.5  170695  9  AC016134                  AC016134 Homo sapi
c  39      69.8        2.5   98951  2  AC090069                  AC090069 Homo sapi
   40      69.8        2.5  208220  9  AC009107                  AC009107 Homo sapi
c  41      68.8        2.5   34744  9  AL589842                  AL589842 Human DNA
c  42      68.6        2.5   38399  9  AY341029                  AY341029 Homo sapi
c  43      68.6        2.5  115396  9  AC010366                  AC010366 Homo sapi
c  44      68.6        2.5  130530  9  AF131217                  AF131217 Homo sapi
c  45      68.6        2.5  168991  2  AC025462                  AC025462 Homo sapi

                                  ALIGNMENTS

RESULT 1
AX535073
LOCUS       AX535073                2754 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 100 from Patent WO02068633.
ACCESSION   AX535073
VERSION     AX535073.1  GI:25261753
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Macina,R.A., Recipon,H., Chen,S.Y., Sun,Y. and Liu,C.
  TITLE     Compositions and methods relating to lung specific genes and
            proteins
  JOURNAL   Patent: WO 02068633-A 100 06-SEP-2002;
            Diadexus, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2754
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 2754;  DB 6;  Length 2754;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2754;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60

Qy         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120

Qy        121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180

Qy        181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240

Qy      241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300

Qy      301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360

Qy      361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420

Qy      421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480

Qy      481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540

Qy      541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600

Qy      601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660

Qy      661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720

Qy      721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGAACCTGTCCTCATGCCCTTCGTCCT 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGAACCTGTCCTCATGCCCTTCGTCCT 780

Qy      781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840

Qy      841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900

Qy      901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960

Qy      961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020

Qy     1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080

Qy     1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140

Qy     1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200

Qy     1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260

Qy     1261 CTCAGACAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 CTCAGACAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320

Qy     1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380

Qy     1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440

Qy     1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500

Qy     1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560

Qy     1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620

Qy     1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680

Qy     1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740

Qy     1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800

Qy     1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860

Qy     1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920

Qy     1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980

Qy     1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040

Qy     2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100

Qy     2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160

Qy     2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220

Qy     2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280

Qy     2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340

Qy     2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400
```

```
Qy        2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460

Qy        2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520

Qy        2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580

Qy        2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640

Qy        2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700

Qy        2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 TGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754


RESULT 2
AC079988/c
LOCUS       AC079988               211305 bp    DNA     linear   PRI 09-MAR-2002
DEFINITION  Homo sapiens BAC clone RP11-795C1 from 2, complete sequence.
ACCESSION   AC079988
VERSION     AC079988.6  GI:18873965
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 211305)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 211305)
  AUTHORS   Waligorski,J., Haakenson,W. and Boyer,E.
  TITLE     The sequence of Homo sapiens BAC clone RP11-795C1
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 211305)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-SEP-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 211305)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-FEB-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 211305)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Feb 23, 2002 this sequence version replaced gi:18042517.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0795C01
            --------------.
```

```
            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO. For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-60M20, 2000 bp overlap.
            Actual start of this clone is at base position 1 of RP11-795C1;
            actual end is at base position 111002 of RP11-6906.

            Single plasmid region exists between 76851 and 76869. Polymorphisms
            exists between AC067960, AC013399 and AC079988. Data from AC067960
            and AC013399 was used to finish AC079988.
FEATURES             Location/Qualifiers
     source          1..211305
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-795C1"
                     /clone_lib="RPCI-11"
     repeat_region   1..115
                     /rpt_family="MIR"
     repeat_region   1042..1153
                     /rpt_family="MER1_type"
     repeat_region   1682..1832
                     /rpt_family="(TG)n"
     repeat_region   1974..2095
                     /rpt_family="Alu"
     repeat_region   2096..2144
                     /rpt_family="(TAAA)n"
     repeat_region   2145..2301
                     /rpt_family="Alu"
     repeat_region   2371..2472
                     /rpt_family="MIR"
     repeat_region   2491..2583
                     /rpt_family="MIR"
     repeat_region   3648..3699
                     /rpt_family="T-rich"
     repeat_region   3726..3838
                     /rpt_family="MIR"
     repeat_region   4074..4169
                     /rpt_family="MIR"
     repeat_region   4375..4466
                     /rpt_family="MER1_type?"
     repeat_region   4627..4725
                     /rpt_family="L2"
     repeat_region   4825..4920
```

```
                     /rpt_family="L2"
     repeat_region   4921..5082
                     /rpt_family="(TG)n"
     repeat_region   5111..5368
                     /rpt_family="(TG)n"
     repeat_region   5429..5655
                     /rpt_family="(TG)n"
     repeat_region   6114..6269
                     /rpt_family="MIR"
     repeat_region   6325..6431
                     /rpt_family="MIR"
     repeat_region   6541..6674
                     /rpt_family="MIR"
     repeat_region   6713..6936
                     /rpt_family="MaLR"
     repeat_region   6937..7459
                     /rpt_family="ERV1"
     repeat_region   7460..7663
                     /rpt_family="MaLR"
     repeat_region   7679..7869
                     /rpt_family="MIR"
     repeat_region   7892..8092
                     /rpt_family="MaLR"
     repeat_region   8693..8775
                     /rpt_family="MIR"
     repeat_region   8821..8897
                     /rpt_family="MIR"
     repeat_region   9061..9445
                     /rpt_family="L2"
     repeat_region   9608..9721
                     /rpt_family="(TA)n"
     repeat_region   9723..9772
                     /rpt_family="(GA)n"
     repeat_region   9904..10068
                     /rpt_family="MER1_type"
     repeat_region   10147..10247
                     /rpt_family="L2"
     repeat_region   10496..10921
                     /rpt_family="MaLR"
     repeat_region   11078..11199
                     /rpt_family="MIR"
     repeat_region   11298..11841
                     /rpt_family="MaLR"
     repeat_region   12041..12220
                     /rpt_family="MIR"
     repeat_region   12379..12465
                     /rpt_family="MIR"
     repeat_region   12543..13210
                     /rpt_family="L1"
     repeat_region   14121..14169
                     /rpt_family="L2"
     repeat_region   14163..14729
                     /rpt_family="L2"
     repeat_region   14802..14966
                     /rpt_family="MER1_type"
     repeat_region   14974..15129
                     /rpt_family="L2"
     repeat_region   15286..15477
                     /rpt_family="ERVL"
     repeat_region   15518..15888
                     /rpt_family="L1"
     repeat_region   15889..16196
                     /rpt_family="Alu"
     repeat_region   16197..16848
                     /rpt_family="L1"
     repeat_region   16861..16884
                     /rpt_family="AT_rich"
     repeat_region   17160..17297
                     /rpt_family="L1"
     repeat_region   17298..17591
                     /rpt_family="Alu"
     repeat_region   17592..18212
                     /rpt_family="L1"
```

```
     repeat_region   18213..18522
                     /rpt_family="Alu"
     repeat_region   18523..19061
                     /rpt_family="L1"
     repeat_region   19062..19198
                     /rpt_family="ERVL"
     repeat_region   19199..19512
                     /rpt_family="Alu"

  Query Match             98.7%;  Score 2718.4;  DB 9;  Length 211305;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2722;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      90196 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 90137

Qy         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      90136 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 90077

Qy        121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180
              |||||||||| |||||||||||||||||||||||||||||||| | ||||||||| | | 
Db      90076 GAGTGGATGAAGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAGTGGCACTGGGCAAGGC 90017

Qy        181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      90016 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 89957

Qy        241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89956 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 89897

Qy        301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89896 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 89837

Qy        361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89836 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 89777

Qy        421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89776 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 89717

Qy        481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89716 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 89657

Qy        541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89656 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 89597

Qy        601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89596 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 89537

Qy        661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89536 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 89477

Qy        721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89476 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 89417

Qy        781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89416 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 89357

Qy        841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      89356 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 89297
```

```
Qy      901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    89296 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 89237

Qy      961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    89236 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 89177

Qy     1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    89176 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 89117

Qy     1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    89116 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 89057

Qy     1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    89056 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 88997

Qy     1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88996 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 88937

Qy     1261 CTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88936 CTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 88877

Qy     1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88876 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 88817

Qy     1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88816 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 88757

Qy     1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88756 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 88697

Qy     1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88696 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 88637

Qy     1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88636 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 88577

Qy     1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88576 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 88517

Qy     1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88516 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 88457

Qy     1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88456 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 88397

Qy     1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88396 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 88337

Qy     1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88336 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 88277

Qy     1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88276 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 88217

Qy     1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88216 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 88157

Qy     2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88156 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 88097

Qy     2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88096 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 88037

Qy     2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    88036 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 87977

Qy     2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87976 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 87917

Qy     2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87916 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 87857

Qy     2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87856 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 87797

Qy     2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87796 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 87737

Qy     2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGAGTATCTGTC 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87736 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGAGTATCTGTC 87677

Qy     2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87676 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 87617

Qy     2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87616 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 87557

Qy     2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    87556 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 87497

Qy     2701 TGCAAATAAAACTCACATTTGTAAGAAA 2728
            ||||||||||||||||||||||||||||
Db    87496 TGCAAATAAAACTCACATTTGTAAGAAA 87469


RESULT 3
AC140107/c
LOCUS       AC140107                225203 bp    DNA     linear   HTG 21-FEB-2003
DEFINITION  Homo sapiens chromosome 17 clone RP11-1201C10 map 17, WORKING DRAFT
            SEQUENCE, 4 unordered pieces.
ACCESSION   AC140107
VERSION     AC140107.1  GI:28460978
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 225203)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-1201C10
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 225203)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
```

```
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L29363
                Center clone name: 1201_C_10
            -------------- Summary Statistics
                Sequencing vector: Plasmid; n/a; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 223547 bases at least Q40
                Consensus quality: 224174 bases at least Q30
                Consensus quality: 224530 bases at least Q20
                Insert size: 221000; agarose-fp
                Insert size: 224903; sum-of-contigs
                Quality coverage: 14.9 in Q20 bases; agarose-fp
                Quality coverage: 14.6 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    46920: contig of 46920 bp in length
            *    46921    47020: gap of 100 bp
            *    47021    51090: contig of 4070 bp in length
            *    51091    51190: gap of 100 bp
            *    51191   163725: contig of 112535 bp in length
            *   163726   163825: gap of 100 bp
            *   163826   225203: contig of 61378 bp in length.
FEATURES             Location/Qualifiers
     source          1. .225203
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-1201C10"
                     /clone_lib="RPCI-11 Male BAC segment 5"
     misc_feature    1. .46920
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    47021. .51090
                     /note="assembly_fragment"
     misc_feature    51191. .163725
```

```
                     /note="assembly_fragment"
     misc_feature    163826. .225203
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             98.7%;  Score 2718.4;  DB 2;  Length 225203;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 2722;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132076 GCCAGAAGCAGCCTCAGCTTGGCAAGGTGTGGAGATGACTGCTGTTCCCTTCGCATTTGG 132017

Qy         61 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     132016 GGAAAACAGGCTCCCTCGGTAGCTCGATGATCCTCTTTTGATCTTGTGTGACCTCCTGGA 131957

Qy        121 GAGTGGATGACGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAATTGCACTGGGCGATGT 180
              |||||||||| |||||||||||||||||||||||||||||||| | ||||||||||| |
Db     131956 GAGTGGATGAAGCTGGTGGCCTTAGCTTTTCTAGACAGTGTAAGTGGCACTGGGCGAGGC 131897

Qy        181 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131896 CCCCAGAGCAGGGCAAGGTCTCTAGAGCGGGTCTCCCACATGACTGGCTTCACACAGGCA 131837

Qy        241 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131836 CTTCCGCTCGGGTTGCATGCTCTGTGTCATCTTACCGGTCCAGGGTTGCAGGTAGGAAAT 131777

Qy        301 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131776 GTTTGTACCCTCTTCTGATTGCCACCTCCTTCCCATCGCCCCTTAGGGACAGGGCTTGAG 131717

Qy        361 GGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 420
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     131716 GGCCAGTGAGGTGCTGGTCAGGCACCCCAGGCCTCCTTGGGACCTGCCCAGGGGCACCCT 131657

Qy        421 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131656 GAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTT 131597

Qy        481 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131596 TCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAG 131537

Qy        541 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131536 TGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTG 131477

Qy        601 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131476 TCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGA 131417

Qy        661 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131416 TAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAA 131357

Qy        721 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131356 CCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCT 131297

Qy        781 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131296 GAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGA 131237

Qy        841 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131236 ACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCC 131177

Qy        901 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     131176 CTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTA 131117

Qy        961 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131116 TGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAA 131057

Qy       1021 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     131056 AATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTG 130997

Qy       1081 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130996 GATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGAT 130937

Qy       1141 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130936 TTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTC 130877

Qy       1201 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130876 TGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGC 130817

Qy       1261 CTCAGACAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130816 CTCAGACAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGG 130757

Qy       1321 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130756 GGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTG 130697

Qy       1381 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130696 TGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCA 130637

Qy       1441 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130636 CCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGT 130577

Qy       1501 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130576 TCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATG 130517

Qy       1561 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130516 CGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCT 130457

Qy       1621 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130456 TGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGGACCTT 130397

Qy       1681 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130396 TTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGCTACAA 130337

Qy       1741 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130336 AGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGCCCTCC 130277

Qy       1801 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130276 ATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCCTCAAA 130217

Qy       1861 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130216 TGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCCCTGTG 130157

Qy       1921 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130156 GCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGCCACTG 130097

Qy       1981 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130096 CAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAGGCTTT 130037

Qy       2041 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     130036 GTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAGGCCCC 129977

Qy       2101 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129976 CACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCCACAGC 129917

Qy       2161 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129916 TACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTGCAGAT 129857

Qy       2221 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129856 CTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCCTGAGG 129797

Qy       2281 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129796 CCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAAC 129737

Qy       2341 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129736 CACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAA 129677

Qy       2401 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129676 GCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTA 129617

Qy       2461 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129616 ACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTC 129557

Qy       2521 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129556 GGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGT 129497

Qy       2581 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129496 TGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACT 129437

Qy       2641 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129436 TGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAA 129377

Qy       2701 TGCAAATAAAACTCACATTTGTAAGAAA 2728
              ||||||||||||||||||||||||||||
Db     129376 TGCAAATAAAACTCACATTTGTAAGAAA 129349


RESULT 4
AX535064/c
LOCUS       AX535064                2399 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 91 from Patent WO02068633.
ACCESSION   AX535064
VERSION     AX535064.1  GI:25261735
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Macina,R.A., Recipon,H., Chen,S.Y., Sun,Y. and Liu,C.
  TITLE     Compositions and methods relating to lung specific genes and
            proteins
  JOURNAL   Patent: WO 02068633-A 91 06-SEP-2002;
            Diadexus, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .2399
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             87.1%;  Score 2397.4;  DB 6;  Length 2399;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 2398;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        356 TTGAGGGCCAGTGAGGCGCTGGTCAGGCACCCCAGGCCTCCTTGGACCTGCCCAGGGGC 415
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       2399 TTGAGGGCCAGTGAGGTGCTGGTCAGGCACCCCAGGCCTCCTTGGACCTGCCCAGGGGC 2340

Qy        416 ACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCT 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2339 ACCCTGAGAGCTCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCT 2280

Qy        476 GGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGT 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2279 GGCTTTCCTCCCTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGT 2220

Qy        536 GCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGG 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2219 GCTAGTGGCCAGGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGG 2160

Qy        596 CAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGA 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2159 CAGTGTCCTTAGGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGA 2100

Qy        656 TGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGG 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2099 TGGGATAAGGTGGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGG 2040

Qy        716 GGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTC 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2039 GGAAACCAAACCATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTC 1980

Qy        776 GTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAA 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1979 GTCCTGAGGAGCCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAA 1920

Qy        836 AGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGA 895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1919 AGAGAACAACATCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGA 1860

Qy        896 CTGCCCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTG 955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1859 CTGCCCTTGTCAGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTG 1800

Qy        956 ATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAA 1015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1799 ATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAA 1740

Qy       1016 AAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTA 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1739 AAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTA 1680

Qy       1076 TTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATA 1135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1679 TTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATA 1620

Qy       1136 TTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAA 1195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1619 TTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAA 1560

Qy       1196 AAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAG 1255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1559 AAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAG 1500

Qy       1256 AAAGCCTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAA 1315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1499 AAAGCCTCAGACAGTAAGGGCTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAA 1440

Qy       1316 GCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTG 1375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1439 GCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTG 1380

Qy       1376 TCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGA 1435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1379 TCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGA 1320

Qy       1436 GTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTT 1495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1319 GTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTT 1260

Qy       1496 CCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGT 1555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1259 CCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGT 1200

Qy       1556 TGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCAC 1615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1199 TGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCAC 1140

Qy       1616 ACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGG 1675
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1139 ACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTGAGG 1080

Qy       1676 ACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGC 1735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1079 ACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCCTGC 1020

Qy       1736 TACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGC 1795
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1019 TACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTTGAGGGGCCGCCCCCCGGC 960

Qy       1796 CCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCC 1855
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        959 CCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCCATTGGAGGATGGACAGCC 900

Qy       1856 TCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCC 1915
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        899 TCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGGCTGTGGCCATCCAGCCCC 840

Qy       1916 CTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGC 1975
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        839 CTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACTCCAGGGGCAGACAGCAGC 780

Qy       1976 CACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAG 2035
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 CACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAGCTGGGGCTAACAGGCTAG 720

Qy       2036 GCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAG 2095
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 GCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCTCCCTAAAGCAATGGGGAG 660

Qy       2096 GCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCC 2155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 GCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTAGATGGGGGCCAGGGTCCC 600

Qy       2156 ACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTG 2215
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 ACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGTCCAGTCCGGATCCCACTTG 540

Qy       2216 CAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCC 2275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 CAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTGTCACAGTGCTGGCTCTGTCC 480

Qy       2276 TGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGG 2335
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGG 420

Qy       2336 TTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGG 2395
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGG 360

Qy       2396 TGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAG 2455
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAATGATTTCTTTCTAAAG 300

Qy       2456 ATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTAT 2515
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         299 ATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAAAGGTGGTGGGAGTAT 240

Qy        2516 CTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTG 2575
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         239 CTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTG 180

Qy        2576 GGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTT 2635
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         179 GGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCTGAAATAAAGCATATTTT 120

Qy        2636 GCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGATCTGTGCTTGGGT 2695
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         119 GCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTGTATGTGGATCTGTGCTTGGGT 60

Qy        2696 TAGAATGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAAGATGCGGCC 2754
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          59 TAGAATGCAAATAAAACTCACATTTGTAAGAAAAAAAAAAAAAAATAAAAAGATGCGGCC 1


RESULT 5
AR204690
LOCUS       AR204690                 1853 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 3 from patent US 6368794.
ACCESSION   AR204690
VERSION     AR204690.1  GI:21502081
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1853)
  AUTHORS   Daniel,S., Gilmore,J., Stuart,S.G. and Stuve,L.L.
  TITLE     Detection of altered expression of genes regulating cell
            proliferation
  JOURNAL   Patent: US 6368794-A 3 09-APR-2002;
FEATURES             Location/Qualifiers
     source          1..1853
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match            63.5%;  Score 1747.6;  DB 6;  Length 1853;
  Best Local Similarity  99.8%;  Pred. No. 0;
  Matches 1750;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         994 TGTTTTCTCCCCCATCACAAAAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAA 1053
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 TGTTTTCTCCCCCATCACAAAAAAAAAAAATTCTTATTTTTAGTAGACATGTATTTACCAA 60

Qy        1054 AAATATGTACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGT 1113
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 AAATATGTACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGT 120

Qy        1114 TTGCTCTTACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTT 1173
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TTGCTCTTACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAAATATTT 180

Qy        1174 ACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCC 1233
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 ACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCC 240

Qy        1234 ATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTTTTAAAAGAATAAA 1293
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTTTTAAAAGAATAAA 300

Qy        1294 ATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAAC 1353
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAAC 360

Qy        1354 CCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGG 1413
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGG 420

Qy        1414 TGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTG 1473
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 TGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTG 480

Qy        1474 CTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATC 1533
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 CTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATC 540

Qy        1534 CCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGC 1593
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGC 600

Qy        1594 AGTGGCCTCCTGGCCAGCTCACACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCA 1653
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AGTGGCCTCCTGGCCAGCTCACACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCA 660

Qy        1654 CCTAGTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCA 1713
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 CCTAGTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCA 720

Qy        1714 TGTGTCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGG 1773
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 TGTGTCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGG 780

Qy        1774 TTTTGAGGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATC 1833
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 TTTTGAGGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATC 840

Qy        1834 TGCCATTGGAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGT 1893
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 TGCCATTGGAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGT 900

Qy        1894 CCGGCTGTGGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTT 1953
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         901 CCGGCTGTGGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTT 960

Qy        1954 CACTCCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATA 2013
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 CACTCCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATA 1020

Qy        2014 GCAGCTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGA 2073
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 GCAGCTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGA 1080

Qy        2074 TCCTCCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAG 2133
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 TCCTCCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAG 1140

Qy        2134 GTTAGATGGGGGCCAGGGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGG 2193
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 GTTAGATGGGGGCCAGGGTCCCACAGCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGG 1200

Qy        2194 TGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTT 2253
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 TGTCCAGTCCGGATCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTT 1260

Qy        2254 TGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAA 2313
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 TGTCACAGTGCTGGCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTCTGCTGGGAA 1320

Qy        2314 AAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTC 2373
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 AAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGACTGTTTGGGGTGCCTC 1380

Qy        2374 TCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTAT 2433
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 TCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTAT 1440

Qy        2434 TTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCT 2493
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 TTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCT 1500

Qy        2494 TTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGG 2553
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1501 TTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGG 1560

Qy          2554 TGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTG 2613
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1561 TGTGAGAGGTCCTGGGGAGGTGGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTG 1620

Qy          2614 TACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTG 2673
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1621 TACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTGCGGACGAGAAGTCTG 1680

Qy          2674 TATGTGGGATCTGTGCTTGGGTTAGAATGCAAATAAAACTCACATTTGTAAGAAAAAAAA 2733
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db          1681 TATGTGGGATCTGTGCTTGGGTTAGAATGCAAATAAAACTCACATTTGTAAGACAAAAAA 1740

Qy          2734 AAAAAATAAAAAGA 2747
                 |||||  ||| |||
Db          1741 AAAAACAAAAGAGA 1754


RESULT 6
AC139359/c
LOCUS       AC139359                88036 bp    DNA     linear   HTG 01-FEB-2003
DEFINITION  Homo sapiens chromosome 17 clone RP11-1162K6 map 17, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC139359
VERSION     AC139359.1  GI:28191551
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 88036)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-1162K6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 88036)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L29391
                Center clone name: 1162_K_6
            --------------.
            * NOTE: This record contains 94 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      856: contig of 856 bp in length
            *      857      956: gap of 100 bp
            *      957     1784: contig of 828 bp in length
            *     1785     1884: gap of 100 bp
            *     1885     2709: contig of 825 bp in length
            *     2710     2809: gap of 100 bp
            *     2810     3605: contig of 796 bp in length
            *     3606     3705: gap of 100 bp
            *     3706     4555: contig of 850 bp in length
            *     4556     4655: gap of 100 bp
            *     4656     5496: contig of 841 bp in length
            *     5497     5596: gap of 100 bp
            *     5597     6438: contig of 842 bp in length
            *     6439     6538: gap of 100 bp
            *     6539     7389: contig of 851 bp in length
            *     7390     7489: gap of 100 bp
            *     7490     8316: contig of 827 bp in length
            *     8317     8416: gap of 100 bp
            *     8417     9255: contig of 839 bp in length
            *     9256     9355: gap of 100 bp
            *     9356    10180: contig of 825 bp in length
            *    10181    10280: gap of 100 bp
            *    10281    11109: contig of 829 bp in length
            *    11110    11209: gap of 100 bp
            *    11210    12059: contig of 850 bp in length
            *    12060    12159: gap of 100 bp
            *    12160    12982: contig of 823 bp in length
            *    12983    13082: gap of 100 bp
            *    13083    13894: contig of 812 bp in length
            *    13895    13994: gap of 100 bp
            *    13995    14795: contig of 801 bp in length
            *    14796    14895: gap of 100 bp
            *    14896    15736: contig of 841 bp in length
            *    15737    15836: gap of 100 bp
            *    15837    16691: contig of 855 bp in length
            *    16692    16791: gap of 100 bp
            *    16792    17640: contig of 849 bp in length
            *    17641    17740: gap of 100 bp
            *    17741    18591: contig of 851 bp in length
            *    18592    18691: gap of 100 bp
            *    18692    19541: contig of 850 bp in length
            *    19542    19641: gap of 100 bp
            *    19642    20489: contig of 848 bp in length
            *    20490    20589: gap of 100 bp
            *    20590    21422: contig of 833 bp in length
            *    21423    21522: gap of 100 bp
            *    21523    22357: contig of 835 bp in length
            *    22358    22457: gap of 100 bp
            *    22458    23292: contig of 835 bp in length
            *    23293    23392: gap of 100 bp
            *    23393    24230: contig of 838 bp in length
            *    24231    24330: gap of 100 bp
            *    24331    25143: contig of 813 bp in length
            *    25144    25243: gap of 100 bp
            *    25244    26055: contig of 812 bp in length
            *    26056    26155: gap of 100 bp
            *    26156    27000: contig of 845 bp in length
            *    27001    27100: gap of 100 bp
            *    27101    27947: contig of 847 bp in length
            *    27948    28047: gap of 100 bp
            *    28048    28896: contig of 849 bp in length
            *    28897    28996: gap of 100 bp
            *    28997    29847: contig of 851 bp in length
```

```
 *     29848     29947: gap of 100 bp
 *     29948     30783: contig of 836 bp in length
 *     30784     30883: gap of 100 bp
 *     30884     31729: contig of 846 bp in length
 *     31730     31829: gap of 100 bp
 *     31830     32675: contig of 846 bp in length
 *     32676     32775: gap of 100 bp
 *     32776     33614: contig of 839 bp in length
 *     33615     33714: gap of 100 bp
 *     33715     34559: contig of 845 bp in length
 *     34560     34659: gap of 100 bp
 *     34660     35489: contig of 830 bp in length
 *     35490     35589: gap of 100 bp
 *     35590     36408: contig of 819 bp in length
 *     36409     36508: gap of 100 bp
 *     36509     37315: contig of 807 bp in length
 *     37316     37415: gap of 100 bp
 *     37416     38262: contig of 847 bp in length
 *     38263     38362: gap of 100 bp
 *     38363     39213: contig of 851 bp in length
 *     39214     39313: gap of 100 bp
 *     39314     40166: contig of 853 bp in length
 *     40167     40266: gap of 100 bp
 *     40267     41123: contig of 857 bp in length
 *     41124     41223: gap of 100 bp
 *     41224     42077: contig of 854 bp in length
 *     42078     42177: gap of 100 bp
 *     42178     43026: contig of 849 bp in length
 *     43027     43126: gap of 100 bp
 *     43127     44011: contig of 885 bp in length
 *     44012     44111: gap of 100 bp
 *     44112     44943: contig of 832 bp in length
 *     44944     45043: gap of 100 bp
 *     45044     45881: contig of 838 bp in length
 *     45882     45981: gap of 100 bp
 *     45982     46811: contig of 830 bp in length
 *     46812     46911: gap of 100 bp
 *     46912     47727: contig of 816 bp in length
 *     47728     47827: gap of 100 bp
 *     47828     48636: contig of 809 bp in length
 *     48637     48736: gap of 100 bp
 *     48737     49571: contig of 835 bp in length
 *     49572     49671: gap of 100 bp
 *     49672     50521: contig of 850 bp in length
 *     50522     50621: gap of 100 bp
 *     50622     51478: contig of 857 bp in length
 *     51479     51578: gap of 100 bp
 *     51579     52428: contig of 850 bp in length
 *     52429     52528: gap of 100 bp
 *     52529     53390: contig of 862 bp in length
 *     53391     53490: gap of 100 bp
 *     53491     54334: contig of 844 bp in length
 *     54335     54434: gap of 100 bp
 *     54435     55273: contig of 839 bp in length
 *     55274     55373: gap of 100 bp
 *     55374     56210: contig of 837 bp in length
 *     56211     56310: gap of 100 bp
 *     56311     57148: contig of 838 bp in length
 *     57149     57248: gap of 100 bp
 *     57249     58079: contig of 831 bp in length
 *     58080     58179: gap of 100 bp
 *     58180     59004: contig of 825 bp in length
 *     59005     59104: gap of 100 bp
 *     59105     59910: contig of 806 bp in length
 *     59911     60010: gap of 100 bp
 *     60011     60839: contig of 829 bp in length
 *     60840     60939: gap of 100 bp
 *     60940     61784: contig of 845 bp in length
 *     61785     61884: gap of 100 bp
 *     61885     62735: contig of 851 bp in length
 *     62736     62835: gap of 100 bp
 *     62836     63686: contig of 851 bp in length
 *     63687     63786: gap of 100 bp
```

```
 *     63787     64628: contig of 842 bp in length
 *     64629     64728: gap of 100 bp
 *     64729     65571: contig of 843 bp in length

 Query Match             29.8%;  Score 820.6;  DB 2;  Length 88036;
 Best Local Similarity   77.9%;  Pred. No. 7.7e-193;
 Matches  838;  Conservative    0;  Mismatches  238;  Indels    0;  Gaps    0;

Qy        427 TCCTGAAACCCCCACTTAGCTTCCAGACCTTTCTGCAAAAGCTCCTCCTGGCTTTCCTCC 486
              |
Db      73127 TNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 73068

Qy        487 CTCCCCCAATCTATGGGTCACAGCTAACAGATCTGAGGGCAACTGCTGTGCTAGTGGCCA 546
                                                          ||||||||||||||||
Db      73067 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGCTGTGCTAGTGGCCA 73008

Qy        547 GGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTGTCCTTA 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73007 GGGCTGCACCTGCCATCCCCGGCTCTGCCACTTTAGGGCCTTCTAGAGGCAGTGTCCTTA 72948

Qy        607 GGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGATAAGGT 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72947 GGAAGTAGCTCTGAGGCATGGGTTTTCTGCTCCTGTGCAGGGCAGCTGATGGGATAAGGT 72888

Qy        667 GGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAACCAAAC 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72887 GGGGAAGGACGGTCAGTGCTTGGGCCCCAGCTGGCCAGCCTGGCGATGGGGAAACCAAAC 72828

Qy        727 CATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCTGAGGAG 786
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72827 CATGTCCCCCAGCGAAGGGCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCTGAGGAG 72768

Qy        787 CCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGAACAACA 846
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72767 CCCTGAGGTGGGCAGCAGGGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGAACAACA 72708

Qy        847 TCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTC 906
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72707 TCCTCAGCTCCGCACCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTC 72648

Qy        907 AGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTATGATAG 966
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72647 AGCTAGCATACGGTGGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTATGATAG 72588

Qy        967 GGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAAAAATTCT 1026
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72587 GGAATATTATCTTTGAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAAAAATTCT 72528

Qy       1027 TATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTGGATTTT 1086
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72527 TATTTTTAGTAGACATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTGGATTTT 72468

Qy       1087 ATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGATTTTGCC 1146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72467 ATCAATTTAAAAATTGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGATTTTGCC 72408

Qy       1147 TCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGA 1206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72407 TCTTGGCTCTGAAAGCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGA 72348

Qy       1207 CTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGA 1266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72347 CTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGA 72288

Qy       1267 CAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGC 1326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      72287 CAGTAAGGGCTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGC 72228

Qy       1327 ATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGC 1386
              |||||||||||||||||||| |||    |  |    |  |    |||| |
Db      72227 ATTCAGATGAGCGGTTTCTGAATTCCANCCCAGGGCGAGATCATCGTTTNNNNNNNNNN 72168

Qy       1387 TGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGC 1446
```

```
Db      72167 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 72108

Qy       1447 CTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTC 1502
                                                  |     | |  |||   |||
Db      72107 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTCTATACTTTATTCTGTTTC 72052


RESULT 7
AC139359
LOCUS       AC139359               88036 bp    DNA     linear   HTG 01-FEB-2003
DEFINITION  Homo sapiens chromosome 17 clone RP11-1162K6 map 17, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC139359
VERSION     AC139359.1  GI:28191551
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 88036)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 17, clone RP11-1162K6
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 88036)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-FEB-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L29391
                Center clone name: 1162_K_6
            --------------.
            * NOTE: This record contains 94 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *           1      856: contig of 856 bp in length
            *         857      956: gap of 100 bp
            *         957     1784: contig of 828 bp in length
            *        1785     1884: gap of 100 bp
            *        1885     2709: contig of 825 bp in length
            *        2710     2809: gap of 100 bp
            *        2810     3605: contig of 796 bp in length
            *        3606     3705: gap of 100 bp
            *        3706     4555: contig of 850 bp in length
            *        4556     4655: gap of 100 bp
            *        4656     5496: contig of 841 bp in length
            *        5497     5596: gap of 100 bp
            *        5597     6438: contig of 842 bp in length
            *        6439     6538: gap of 100 bp
            *        6539     7389: contig of 851 bp in length
            *        7390     7489: gap of 100 bp
            *        7490     8316: contig of 827 bp in length
            *        8317     8416: gap of 100 bp
            *        8417     9255: contig of 839 bp in length
            *        9256     9355: gap of 100 bp
            *        9356    10180: contig of 825 bp in length
            *       10181    10280: gap of 100 bp
            *       10281    11109: contig of 829 bp in length
            *       11110    11209: gap of 100 bp
            *       11210    12059: contig of 850 bp in length
            *       12060    12159: gap of 100 bp
            *       12160    12982: contig of 823 bp in length
            *       12983    13082: gap of 100 bp
            *       13083    13894: contig of 812 bp in length
            *       13895    13994: gap of 100 bp
            *       13995    14795: contig of 801 bp in length
            *       14796    14895: gap of 100 bp
            *       14896    15736: contig of 841 bp in length
            *       15737    15836: gap of 100 bp
            *       15837    16691: contig of 855 bp in length
            *       16692    16791: gap of 100 bp
            *       16792    17640: contig of 849 bp in length
            *       17641    17740: gap of 100 bp
            *       17741    18591: contig of 851 bp in length
            *       18592    18691: gap of 100 bp
            *       18692    19541: contig of 850 bp in length
            *       19542    19641: gap of 100 bp
            *       19642    20489: contig of 848 bp in length
            *       20490    20589: gap of 100 bp
            *       20590    21422: contig of 833 bp in length
            *       21423    21522: gap of 100 bp
            *       21523    22357: contig of 835 bp in length
            *       22358    22457: gap of 100 bp
            *       22458    23292: contig of 835 bp in length
            *       23293    23392: gap of 100 bp
            *       23393    24230: contig of 838 bp in length
            *       24231    24330: gap of 100 bp
            *       24331    25143: contig of 813 bp in length
            *       25144    25243: gap of 100 bp
            *       25244    26055: contig of 812 bp in length
            *       26056    26155: gap of 100 bp
            *       26156    27000: contig of 845 bp in length
            *       27001    27100: gap of 100 bp
            *       27101    27947: contig of 847 bp in length
            *       27948    28047: gap of 100 bp
            *       28048    28896: contig of 849 bp in length
            *       28897    28996: gap of 100 bp
            *       28997    29847: contig of 851 bp in length
            *       29848    29947: gap of 100 bp
            *       29948    30783: contig of 836 bp in length
            *       30784    30883: gap of 100 bp
            *       30884    31729: contig of 846 bp in length
            *       31730    31829: gap of 100 bp
            *       31830    32675: contig of 846 bp in length
            *       32676    32775: gap of 100 bp
            *       32776    33614: contig of 839 bp in length
            *       33615    33714: gap of 100 bp
            *       33715    34559: contig of 845 bp in length
            *       34560    34659: gap of 100 bp
            *       34660    35489: contig of 830 bp in length
```

```
       *     35490     35589: gap of 100 bp
       *     35590     36408: contig of 819 bp in length
       *     36409     36508: gap of 100 bp
       *     36509     37315: contig of 807 bp in length
       *     37316     37415: gap of 100 bp
       *     37416     38262: contig of 847 bp in length
       *     38263     38362: gap of 100 bp
       *     38363     39213: contig of 851 bp in length
       *     39214     39313: gap of 100 bp
       *     39314     40166: contig of 853 bp in length
       *     40167     40266: gap of 100 bp
       *     40267     41123: contig of 857 bp in length
       *     41124     41223: gap of 100 bp
       *     41224     42077: contig of 854 bp in length
       *     42078     42177: gap of 100 bp
       *     42178     43026: contig of 849 bp in length
       *     43027     43126: gap of 100 bp
       *     43127     44011: contig of 885 bp in length
       *     44012     44111: gap of 100 bp
       *     44112     44943: contig of 832 bp in length
       *     44944     45043: gap of 100 bp
       *     45044     45881: contig of 838 bp in length
       *     45882     45981: gap of 100 bp
       *     45982     46811: contig of 830 bp in length
       *     46812     46911: gap of 100 bp
       *     46912     47727: contig of 816 bp in length
       *     47728     47827: gap of 100 bp
       *     47828     48636: contig of 809 bp in length
       *     48637     48736: gap of 100 bp
       *     48737     49571: contig of 835 bp in length
       *     49572     49671: gap of 100 bp
       *     49672     50521: contig of 850 bp in length
       *     50522     50621: gap of 100 bp
       *     50622     51478: contig of 857 bp in length
       *     51479     51578: gap of 100 bp
       *     51579     52428: contig of 850 bp in length
       *     52429     52528: gap of 100 bp
       *     52529     53390: contig of 862 bp in length
       *     53391     53490: gap of 100 bp
       *     53491     54334: contig of 844 bp in length
       *     54335     54434: gap of 100 bp
       *     54435     55273: contig of 839 bp in length
       *     55274     55373: gap of 100 bp
       *     55374     56210: contig of 837 bp in length
       *     56211     56310: gap of 100 bp
       *     56311     57148: contig of 838 bp in length
       *     57149     57248: gap of 100 bp
       *     57249     58079: contig of 831 bp in length
       *     58080     58179: gap of 100 bp
       *     58180     59004: contig of 825 bp in length
       *     59005     59104: gap of 100 bp
       *     59105     59910: contig of 806 bp in length
       *     59911     60010: gap of 100 bp
       *     60011     60839: contig of 829 bp in length
       *     60840     60939: gap of 100 bp
       *     60940     61784: contig of 845 bp in length
       *     61785     61884: gap of 100 bp
       *     61885     62735: contig of 851 bp in length
       *     62736     62835: gap of 100 bp
       *     62836     63686: contig of 851 bp in length
       *     63687     63786: gap of 100 bp
       *     63787     64628: contig of 842 bp in length
       *     64629     64728: gap of 100 bp
       *     64729     65571: contig of 843 bp in length

  Query Match             29.2%;  Score 803;  DB 2;  Length 88036;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-188;
  Matches  803;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        861 CCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTCAGCTAGCATACGGT 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82558 CCCCTCATCCTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTCAGCTAGCATACGGT 82617
```

```
Qy        921 GGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTATGATAGGGAATATTATCTTT 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82618 GGGCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTATGATAGGGAATATTATCTTT 82677

Qy        981 GAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAAAATTCTTATTTTTAGTAGAC 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82678 GAACCCAATGAAGTGTTTTCTCCCCCATCACAAAAAAAAAAATTCTTATTTTTAGTAGAC 82737

Qy       1041 ATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAAT 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82738 ATGTATTTACCAAAAATATGTACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAAT 82797

Qy       1101 TGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAA 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82798 TGTGGAAATTTGTTTGCTCTTACGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAA 82857

Qy       1161 GCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGAC 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82858 GCCCAAAATATTTACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGAC 82917

Qy       1221 CTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTT 1280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82918 CTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTT 82977

Qy       1281 TTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGATGAGCGG 1340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82978 TTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGATGAGCGG 83037

Qy       1341 TTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGCCCCCCT 1400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83038 TTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGCCCCCCT 83097

Qy       1401 TGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGTGGGAAT 1460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83098 TGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGTGGGAAT 83157

Qy       1461 GCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGCTGGGCC 1520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83158 GCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGCTGGGCC 83217

Qy       1521 TTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACTTTGAGC 1580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83218 TTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACTTTGAGC 83277

Qy       1581 AGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCTTGTCCTGGGAGGGGCAGCCT 1640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83278 AGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCTCACACTCTTGTCCTGGGAGGGGCAGCCT 83337

Qy       1641 GATCTCACCTCCACCTAGTACCT 1663
              |||||||||||||||||||||||
Db      83338 GATCTCACCTCCACCTAGTACCT 83360


RESULT 8
AX535072
LOCUS       AX535072                 960 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 99 from Patent WO02068633.
ACCESSION   AX535072
VERSION     AX535072.1  GI:25261751
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Macina,R.A., Recipon,H., Chen,S.Y., Sun,Y. and Liu,C.
  TITLE     Compositions and methods relating to lung specific genes and
            proteins
  JOURNAL   Patent: WO 02068633-A 99 06-SEP-2002;
            Diadexus, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .960
                     /organism="Homo sapiens"
```

```
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"
ORIGIN

  Query Match             22.2%;  Score 612.4;  DB 6;  Length 960;
  Best Local Similarity   99.4%;  Pred. No. 4e-141;
  Matches  646;  Conservative    0;  Mismatches    1;  Indels    3;  Gaps    3;

Qy       1658 GTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTG 1717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 GTACCTTGGGGACTGAGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTG 100

Qy       1718 TCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTT 1777
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 TCCAGCTGCTCTTCCTGCTACAAAGGGGACTGCTCACAGTGGCCTCAGCTTGGTGGTTTT 160

Qy       1778 GAGGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCC 1837
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 GAGGGGCCGCCCCCCGGCCCTCCATAAGGGTATCCTGGGCCTGAGAATTCTGCATCTGCC 220

Qy       1838 ATTGGAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGG 1897
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 ATTGGAGGATGGACAGCCTCAAATGGAAGGAGTCCCACGGGAGATGGGTCCGAGGTCCGG 280

Qy       1898 CTGTGGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACT 1957
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 CTGTGGCCATCCAGCCCCCTGTGGCTTGTCCAGCCTCTGTGCACCCCTGGTGTCTTCACT 340

Qy       1958 CCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAG 2017
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 CCAGGGGCAGACAGCAGCCACTGCAGTTCCTTTCTTCGTGAGTAACAGTAGTGATAGCAG 400

Qy       2018 CTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCT 2077
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 CTGGGGCTAACAGGCTAGGCTTTGTGTTCTGCGCATTTGGTCAGCTTCTCACTCGATCCT 460

Qy       2078 CCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTA 2137
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 CCCTAAAGCAATGGGGAGGCCCCCACTAGCCCAGTTTTCAGGAAGTCAACTGGGAGGTTA 520

Qy       2138 GATGGGGGCCAGGGTCCCACA-GCTACTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGT 2196
              ||||||||||||||||||||| ||| ||||||||||||||||||||||||||||||||||
Db        521 GATGGGGGCCAGGGTCCCACATGCTGCTGATGGCCCGAGCCAGGTTGAGCTTCCTGGTGT 580

Qy       2197 CCAGTCCGGAT-CCCACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTG 2255
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 CCAGTCCGGATCCCCACTTGCAGATCTCATGCTCTCAGATAGGTGGGACAAGTTCTTTTG 640

Qy       2256 TCACAGTGCTGG-CTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTC 2304
              |||||||||||| |||||||||||||||||||||||||||||||||||||
Db        641 TCACAGTGCTGGCCTCTGTCCTGAGGCCTCATTGCTGGCTGGGTGTGCTC 690


RESULT 9
CQ522682
LOCUS       CQ522682                 611 bp    DNA     linear   PAT 30-JAN-2004
DEFINITION  Sequence 54549 from Patent WO0160860.
ACCESSION   CQ522682
VERSION     CQ522682.1  GI:41488946
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schlegel,R., Endege,W.O. and Monahan,J.E.
  TITLE     Genes differentially expressed in human prostate cancer and their
            use
  JOURNAL   Patent: WO 0160860-A 54549 23-AUG-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..611
                     /organism="Homo sapiens"
```

```
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             20.1%;  Score 554;  DB 6;  Length 611;
  Best Local Similarity   99.5%;  Pred. No. 1.4e-126;
  Matches  565;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       1194 AAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAA 1253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         44 AAAAGTCTGCTGACTACTGAGCCAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAA 103

Qy       1254 AGAAAGCCTCAGACAGTAAGGGC-TTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTG 1312
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        104 AGAAAGCCTCAGACAGTAAGGGCTTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTG 163

Qy       1313 GAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTC 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 GAAGCAGGGGAAGCATTCAGATGAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTC 223

Qy       1373 GTGTCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGC 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 GTGTCCTGTGTGGCTGGCGAGCCCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGC 283

Qy       1433 AGAGTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTG 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 AGAGTCCACCCCGCCTCGTGGTGGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTG 343

Qy       1493 CTTCCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAG 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 CTTCCTGTTCCTGGGGACCCGCTGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAG 403

Qy       1553 GGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCT 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 GGTTGATGCGTGGAGAAGGACTTTGAGCAGTGGTGGGCAGCAGTGGCCTCCTGGCCAGCT 463

Qy       1613 CACACTCTTGTCCTGGGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTG 1672
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        464 CACACTCTTGTCCTGNGAGGGGCAGCCTGATCTCACCTCCACCTAGTACCTTGGGGACTG 523

Qy       1673 AGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTCCAGCTGCTCTTCC 1732
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        524 AGGACCTTTTGGCTTCTCTGGAGCCTGCAAGCCTCTTCCCATGTGTNCAGCTGCTCTTCC 583

Qy       1733 TGCTACAAAGGGGACTGCTCACAGTGGC 1760
              ||||||||||||||||||||||||||||
Db        584 TGCTACAAAGGGGACTGCTCACAGTGGC 611


RESULT 10
AX262096
LOCUS       AX262096                 395 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 176 from Patent WO0172781.
ACCESSION   AX262096
VERSION     AX262096.1  GI:16511048
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D.,
            Sudduth-Klinger,J., Reinhard,C., He,Z., Randazzo,F., Kennedy,G.C.,
            Pot,D., Kassam,A., Lamson,G., Drmanac,R., Crkvenjakov,R.,
            Dickson,M., Drmanac,S., Labat,I., Leshkowitz,D., Kita,D.,
            Garcia,V., Jones,L.W. and Stache-Crain,B.
  JOURNAL   Patent: WO 0172781-A 176 04-OCT-2001;
            Chiron Corporation (US) ; Hyseq Inc. (US)
FEATURES             Location/Qualifiers
     source          1..395
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

```
ORIGIN

  Query Match             13.0%;  Score 357.4;  DB 6;  Length 395;
  Best Local Similarity   96.9%;  Pred. No. 1.2e-77;
  Matches  375;  Conservative    0;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       1216 CAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGG 1275
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 CAGACCTCCATTACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGG 68

Qy       1276 C-TTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGAT 1334
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 CTTTTTTTAAAAGAATAAAATGACTTGGTTTGCGCTTGGAAGCAGGGGAAGCATTCAGAT 128

Qy       1335 GAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGC 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 GAGCGGTTTCTGCATTAACCCTGCCTATCACGCATCTCGTGTCCTGTGTGGCTGGCGAGC 188

Qy       1395 CCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGT 1454
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 CCCCCTTGGAAGGTTCTGGTGCTTCAGCTGGCTCCTGCAGAGTCCACCCCGCCTCGTGGT 248

Qy       1455 GGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGC 1514
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 GGGAATGCAGAGCCCTTTGCTTTCCTTCTTGCCGCCTGCTTCCTGTTCCTGGGGACCCGC 308

Qy       1515 TGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGTTGATGCGTGGAGAAGGACT 1574
              |||||||||||||||||||||||||||||||||||||||| ||||||| | |||||||||
Db        309 TGGGCCTTTGGTCTGCATCCCCTGGCCAGGTCCCTCAGGGCTGATGCGCGTAGAAGGACT 368

Qy       1575 TTGAGCAGTGGTGGGCAGCAGTGGCCT 1601
              ||||||||||||||       || |||
Db        369 TTGAGCAGTGGTGGCAGCACTTGCCCT 395


RESULT 11
G38425/c
LOCUS       G38425                   406 bp    DNA     linear   STS 23-JUN-1998
DEFINITION  SHGC-58087 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION   G38425
VERSION     G38425.1  GI:3249199
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Myers,R.M.
  TITLE     Human STSs (1998)
  JOURNAL   Unpublished (1998)
COMMENT
            Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu
            Primer A: AGCTGACAAGGGCAGTCACA
            Primer B: GGGGAAGTTTTCAGGCCTTC
            STS size: 102
            PCR Profile:
                 Initial incubation: 95 degrees C for 10 minutes

                 Denaturation:       94 degrees C for 30 seconds
                 Annealing:          60 degrees C for 30 seconds
                 Polymerization:     72 degrees C for 23 seconds
                 PCR Cycles:         30
                 Thermal Cycler:     Perkin Elmer 9700
            Protocol:
                 Template:               25 ng
                 Primer:                 each 1 uM
                 dNTPs:                  each 200 uM
```

```
                 AmpliTaq Gold Polymerase: 0.07 units/ul
                 Total Vol:                5 ul

            Buffer:
                 MgCl2:          2.5 mM
                 KCl:           50 mM
                 Tris-HCl:      10 mM
                 pH:             8.3

            Prepared with primer pairs derived from W86076 -- Unigene.
FEATURES             Location/Qualifiers
     source          1. .406
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="2"
                     /clone_lib="Human"
     STS             100. .201
     primer_bind     100. .119
     primer_bind     complement(182. .201)
ORIGIN

  Query Match              9.5%;  Score 261.2;  DB 11;  Length 406;
  Best Local Similarity   89.2%;  Pred. No. 1.1e-53;
  Matches  347;  Conservative    0;  Mismatches   33;  Indels    9;  Gaps    6;

Qy        630 TTTCTGCTCCTGTGCAGGGCAGCTGATGGG-ATAAGGTGGGGAAGGACGG--TCAGTGCT 686
              || || |   ||    || ||||||||||| |||||||||||||||||||   |||||||
Db        387 TTNCTCCCTGTGCAAGGGCCAGCTGATGGGNATAAGGTGGGGAAGGACGGGTCCAGTGCT 328

Qy        687 TGGGCCCCAGCTGGCC--AGCCTGGCGATGGGGAAACCAAACCATGTCCCCCAGCGAAGG 744
              ||| ||| ||||| ||    |  |  ||||||||||||||||||||||||||||||||||
Db        327 TGGCCCCNAGCTGNCCCAGCCTGGCGGATGGGGAAACCAAACCATGTCCCCCAGCGAAGG 268

Qy        745 GCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCTGAGGAGCCCTGAGGTGGGCAGCAG 804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GCCAGAGTGGGAACCTGTCCTCATGCCCTTCGTCCTGAGGAGCCCTGAGGTGGGCAGCAG 208

Qy        805 GGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGAACAACATCC-TCAGCTCCGCACCC 863
              ||||||||||||||||||||||||||||||||||||||||||||| ||   |||||||||
Db        207 GGGCCAGGGGAAGTTTTCAGGCCTTCATCAAAGAGAACAACATCCGTCGAGTCCGCACCC 148

Qy        864 CTC-ATCCTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTCAGCTAGCATACGGTGG 922
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||     ||
Db        147 CTCGATCCTGTATCAGCACTTACCGGTGTGTGACTGCCCTTGTCAGCTAGCAT--CACGG 90

Qy        923 GCCCACCTGGCCCACTGGCTGTTTATGCCACTGATTTATGATAGGGAATATTATCTTTGA 982
                ||  |  | |||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 NGCCCACCTGNCCACTGGCTGTTTATGCCACTGATTTATGATAGGGAATATTATCTTTGA 30

Qy        983 ACCCAATGAAGTGTTTTCTCCCCCATCAC 1011
              |||||||||||||||||||||||||||||
Db         29 ACCCAATGAAGTGTTTTCTCCCCCATCAC 1


RESULT 12
AX535063/c
LOCUS       AX535063                 251 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 90 from Patent WO02068633.
ACCESSION   AX535063
VERSION     AX535063.1  GI:25261733
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Macina,R.A., Recipon,H., Chen,S.Y., Sun,Y. and Liu,C.
  TITLE     Compositions and methods relating to lung specific genes and
            proteins
  JOURNAL   Patent: WO 02068633-A 90 06-SEP-2002;
            Diadexus, Inc. (US)
FEATURES             Location/Qualifiers
```

```
     source          1..251
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match              8.1%;  Score 224.4;  DB 6;  Length 251;
  Best Local Similarity   98.8%;  Pred. No. 1.6e-44;
  Matches  237;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy       2381 CCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAAT 2440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 CCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAATGTTCATGTATTTAAAAT 192

Qy       2441 GATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAA 2500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATTGGGTGACTCTTTTCTAA 132

Qy       2501 AGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGA 2560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 AGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGGGTCAGGTGGGTGTGAGA 72

Qy       2561 GGTCCTGGGGAGGT--GGGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCT 2618
              ||||||||||||||  | ||||||||||||||||||||||||||||||||||||||||||
Db         71 GGTCCTGGGGAGGTGCGCGCGTTGAGCTCAAAGTTGTCCTACTGCCATGTTTTTGTACCT 12


RESULT 13
AC134136/c
LOCUS       AC134136               190379 bp    DNA     linear   HTG 20-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-298J1, *** SEQUENCING IN PROGRESS
            ***, 2 unordered pieces.
ACCESSION   AC134136
VERSION     AC134136.2  GI:25138971
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 190379)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 190379)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 190379)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 20, 2002 this sequence version replaced gi:23269326.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: KCCX
                Center clone name: CH230-298J1
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 168342 bases at least Q40
                Consensus quality: 171971 bases at least Q30
                Consensus quality: 174193 bases at least Q20
                Estimated insert size: 172482; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    14191: contig of 14191 bp in length
            *    14192    14291: gap of unknown length
            *    14292   190379: contig of 176088 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..190379
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-298J1"
     misc_feature    1..1314
                     /note="wgs_contig"
     misc_feature    10470..12661
                     /note="wgs_contig"
     misc_feature    12712..14191
                     /note="wgs_contig"
     misc_feature    113270..115130
                     /note="wgs_contig"
     misc_feature    complement(170214..171114)
                     /note="clone_boundary
                     clone_end:Sp6
                     site:
                     end_sequence:BZ129373"
     misc_feature    186833..188023
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    189278..190379
                     /note="wgs_end_extension
                     clone_end:Sp6"
ORIGIN

  Query Match             3.2%;  Score 87.6;  DB 2;  Length 190379;
  Best Local Similarity   65.6%;  Pred. No. 2.9e-10;
  Matches  269;  Conservative    0;  Mismatches   99;  Indels   42;  Gaps    8;

Qy       2304 CTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGACTGTTT 2363
              |||||||| | ||||| | |||||     | ||     | | |||  |||| ||||||||
Db       5492 CTGCTGGGGAGAGCTTAGTGGGGC-----TAGTATCAGAGACAAGCAAAGGAGACTGTTT 5438

Qy       2364 GGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAAT 2423
                 |   | ||       |||  | ||||||||  | ||  |||||||||||||| | ||
Db       5437 CATG---CACTGACTTATCTCTTCCTGCTGGGTAAAGGCTTGGTTACTGTGTTCTTTCAT 5381

Qy       2424 GTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATT 2483
              | |      |      |||  | ||   | |        |    |  ||| ||| || |
Db       5380 GCTTCGAAGTGATTCCTGAGGTTTTCTCACA--------CCTACCTTCTTCCTCTAGGTG 5329

Qy       2484 GGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGG 2543
              |||| | ||||||||||||||||| || | || | | ||||||||          ||| |
Db       5328 GGGTAATTCTTTTCTAAAGGTGGTAGGGG-ATTTCTGGGGGTGGTCCA-------GATTG 5277

Qy       2544 GTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGTT-GAGCTCAAAGTTGTCCTACTG 2602
              ||||  || ||||   ||||| |||||| | ||||  | ||   | |||||||| || ||
Db       5276 GTCAAATGTGTGTAGCAGGTCATGGGGAAGGGGGCCATGGAAACCCAAGTTGTCATAATG 5217

Qy       2603 CCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTGCGGA 2662
              ||||||||||||||||||||||||||||||||||||||||| |                |
Db       5216 CCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTAA----------------A 5173

Qy       2663 CGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAATGCAAATAAAAC 2712
               ||||| |||||||||||| ||||||| ||||| ||||||||||||||||
Db       5172 TGAGAAATCTGTATGTGGGCTCTGTGC-TGGGTCAGAATGCAAATAAAAC 5124


RESULT 14
AC136817
LOCUS       AC136817               231580 bp    DNA     linear   HTG 23-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-43B4, *** SEQUENCING IN PROGRESS ***,
            2 unordered pieces.
ACCESSION   AC136817
VERSION     AC136817.2  GI:25188334
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
```

```
REFERENCE   1  (bases 1 to 231580)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 231580)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 231580)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 23, 2002 this sequence version replaced gi:24796690.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
```

```
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: KCLR
                Center clone name: CH230-43B4
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 204259 bases at least Q40
                Consensus quality: 209008 bases at least Q30
                Consensus quality: 211699 bases at least Q20
                Estimated insert size: 211394; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *          1   230160: contig of 230160 bp in length
            *     230161   230260: gap of unknown length
            *     230261   231580: contig of 1320 bp in length.
FEATURES             Location/Qualifiers
     source          1..231580
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-43B4"
     misc_feature    1..1287
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    1338..2663
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    9136..9761
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH263051"
     misc_feature    43876..45375
                     /note="wgs_contig"
     misc_feature    complement(225032..225921)
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH263050"
     misc_feature    229121..230160
                     /note="wgs_end_extension
                     clone_end:T7"
ORIGIN

  Query Match             3.2%;  Score 87.6;  DB 2;  Length 231580;
  Best Local Similarity   65.6%;  Pred. No. 2.9e-10;
  Matches  269;  Conservative    0;  Mismatches   99;  Indels   42;  Gaps    8;

Qy       2304 CTGCTGGGAAAAGCTTTGCGGGGCTTGCTTGGTTAACCACAGAAGAGAAGGGGACTGTTT 2363
              |||||||| | ||||| | |||||     | ||     | | |||  |||| ||||||||
Db     225296 CTGCTGGGGAGAGCTTAGTGGGGC-----TAGTATCAGAGACAAGCAAAGGAGACTGTTT 225350

Qy       2364 GGGGTGCCTCTCTGCAGCCTCCCCGTGCTGGGTGGAAGCACGGTTACTGTGTTCTCTAAT 2423
                 |   | ||       |||  | ||||||||  | ||  |||||||||||||| | ||
Db     225351 CATG---CACTGACTTATCTCTTCCTGCTGGGTAAAGGCTTGGTTACTGTGTTCTTTCAT 225407
```

```
Qy       2424 GTTCATGTATTTAAAATGATTTCTTTCTAAAGATGTAACCTCCACACCTTTCTCCAGATT 2483
              | |      |      |||  | ||   | |        |    |  ||| ||| || |
Db     225408 GCTTCGAAGTGATTCCTGAGGTTTTCTCACA--------CCTACCTTCTTCCTCTAGGTG 225459

Qy       2484 GGGTGACTCTTTTCTAAAGGTGGTGGGAGTATCTGTCGGGGTGGTGTGGCCCTTGGATGG 2543
              |||| | ||||||||||||||||| || | || | | ||||||||          ||| |
Db     225460 GGGTAATTCTTTTCTAAAGGTGGTAGGGG-ATTTCTGGGGGTGGTCCA-------GATTG 225511

Qy       2544 GTCAGGTGGGTGTGAGAGGTCCTGGGGAGGTGGGCGTT-GAGCTCAAAGTTGTCCTACTG 2602
              ||||  || ||||   ||||| |||||| | ||||  | ||   | |||||||| || ||
Db     225512 GTCAAATGTGTGTAGCAGGTCATGGGGAAGGGGGCCATGGAAACCCAAGTTGTCATAATG 225571

Qy       2603 CCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTTACTGTACCATAGTGCGGA 2662
              ||||||||||||||||||||||||||||||||||||||||| |                |
Db     225572 CCATGTTTTTGTACCTGAAATAAAGCATATTTTGCACTTGTAA----------------A 225615

Qy       2663 CGAGAAGTCTGTATGTGGGATCTGTGCTTGGGTTAGAATGCAAATAAAAC 2712
               ||||| |||||||||||| ||||||| ||||| ||||||||||||||||
Db     225616 TGAGAAATCTGTATGTGGGCTCTGTGC-TGGGTCAGAATGCAAATAAAAC 225664


RESULT 15
CNS01DT7
LOCUS       CNS01DT7               182643 bp    DNA     linear   PRI 14-MAY-2001
DEFINITION  Human chromosome 14 DNA sequence BAC R-131H24 of library RPCI-11
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL132642
VERSION     AL132642.4  GI:14041778
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 182643)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 182643)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAY-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On May 15, 2001 this sequence version replaced gi:8217878.
            --------------- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            ---------------
            The following BAC sequence is oriented from the T7 to the SP6 end.
            Upstream BAC (overlapping the T7 end) : R-187A6
            Downstream BAC (overlapping the SP6 end) : R-1089B7 (AC=AL079302)
            --------------- Summary Statistics
            Assembly program: Phrap; version 2.0
            Quality coverage: 7.83x in Q20 bases; sum-of-contigs
            ---------------
            -----------------------
            Overall quality chart :
            Range    :   bases
            0        :
            1  -  9  :
            10 - 19  :
            20 - 29  :      11
            30 - 39  :      83
            40 - 49  :    1023
            50 - 59  :    1882
            60 - 69  :    4058
            70 - 79  :   15227
```

```
               80 - 89  :   59044
               90 - 99  :  101315
               -----------------------
               Percentage of bases with a quality value >= 40 : 99 %.
FEATURES             Location/Qualifiers
     source          1..182643
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="R-131H24"
                     /clone_lib="RPCI-11"
     STS             138601..138720
                     /note="matching EMBL:AA421158
                     RHdb:RH93218
                     dbSTS:STS66221
                     Identified using the e-PCR software (G. Schuler)"
     STS             157690..157856
                     /note="matching EMBL:T03336
                     RHdb:RH53909
                     dbSTS:STS10403
                     Identified using the e-PCR software (G. Schuler)"
     STS             162845..162997
                     /note="matching EMBL:AA610019
                     RHdb:RH94166
                     dbSTS:STS67159
                     Identified using the e-PCR software (G. Schuler)"
     STS             167731..167886
                     /note="matching EMBL:AA195225
                     RHdb:RH65774
                     dbSTS:STS45705
                     Identified using the e-PCR software (G. Schuler)"
ORIGIN

  Query Match             2.9%;  Score 81;  DB 9;  Length 182643;
  Best Local Similarity   54.5%;  Pred. No. 1.3e-08;
  Matches  213;  Conservative    0;  Mismatches  165;  Indels   13;  Gaps    2;

Qy        940 GCTGTTTATGCCACTGATTTATGATAGGGAATATTATCTTTGAACCCAATGAAGTGTTTT 999
              | ||  ||| |   || || ||||| | ||| ||||||||| ||||  |   ||
Db      36767 GATGAATATTCTTTTGTTTGATGATGGAGAACATTATCTTTAAACCACAATTAGGCAAAA 36826

Qy       1000 CTCCCCATCACAAAAAAAAAAAAATTCTTATTTTAGTAGACATGTATTTACCAAAAATAT 1059
                |   |     ||||| ||      || | |       | |  ||| ||  ||||   |
Db      36827 TGCTATCCCATAAAAAAGAATTCCATTTTTCTCATTAGTAGACCTATATAATAAAATATT 36886

Qy       1060 GTACTCAATTATTGTATTTTGGATTTTATCAATTTAAAAATTGTGGAAATTTGTTTGCTC 1119
              ||||| | ||||| | ||||  |||| ||||| | | || || ||||||||||||   ||
Db      36887 GTACTTAGTTATTATGTTTTAAATTTCATCAA-TAATAATTTATGGAAATTTGTTCTGTC 36945

Qy       1120 TTA------------CGCCAACATAATATTGATTTTGCCTCTTGGCTCTGAAAGCCCAAA 1167
              ||             |  ||  |||   |||||| |||||||||||    | |||| |||
Db      36946 TTGTTATGTAAGTGTCTGCATAATATCCTTGATTCTGCCTCTTGGCCTGCACAGCCTAAA 37005

Qy       1168 ATATTTACCGTCTAGCCCGTTACAGAAAAAGTCTGCTGACTACTGAGCCAGACCTCCATT 1227
              ||| ||||  ||| | ||  | ||||| |||| |||||||| |||  |
Db      37006 ATACTTACTATCTGGTCCTCTGCAGAACAAGTATGCTGACTCCTGCTCTTTTTGATGGGA 37065

Qy       1228 ACCTCCATCCCTGTTGGATTATTTAAAGAAAGCCTCAGACAGTAAGGGCTTTTTTAAAAG 1287
                 | ||||  ||     |  | | |||  |      |  || ||      | |||| ||
Db      37066 TATTTCATCTTTGAATTTTGTTCTCAAGCCAAGAAATGGAAGGAATCAGACTGTTAAGAG 37125

Qy       1288 AATAAAATGACTTGGTTTGCGCTTGGAAGCA 1318
                  ||| | | | |   |   |   || ||
Db      37126 CTGGAAAGGCCCTCGGAGGTCTTCTAAACCA 37156

Search completed: November 29, 2004, 11:55:34
Job time : 11599 secs
```